TECHNOLOGIES FOR QUANTITATIVE AND MULTIPLEXED BIOMARKER DETECTION

by

Yunke Song

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

August, 2014

Abstract

Genetic and epigenetic biomarkers have enormous potential to improve healthcare management through their use for prognosis and diagnosis of diseases such as cancer. While DNA and mRNA have been studied for decades, microRNA emerges as a new class of biomarkers that promises to facilitate clinical managements of cancer as well as studies of underling mechanisms governing cancer pathogenesis. My research focuses on the development of new methods and technologies to tackle the critical barriers associated to analyses of DNA and microRNA based biomarkers. This dissertation is devoted for development of (1) genetic mutation detection method that combines excellent specificity of ligation technology and ultrahigh sensitivity of quantum dot (QD) nanosensor-based single molecule detection (SMD), (2) nanomaterial-based method for further enhancement of QD-nanosensor-based assays and (3) quantitative microRNA detection methods that are well customized to provide excellent sensitivity, specificity, multiplexing capability and handling convenience.

Acknowledgement

First and foremost, I would like to thank my advisor Dr. Jeff Tza-Huei Wang for providing me with guidance throughout my PhD program while giving me sufficient freedom to pursue my own research ideas. He has been always a supportive and considerate advisor, helping me demonstrate my best performance in experiments, lab meetings and manuscript preparations. What I learned from Wang Lab would serve as important skill sets for rest of my career.

I am also grateful of having being surrounded by excellent students and postdoctoral fellows in Wang Lab. I am glad to have an opportunity to work closely with Kelvin, from whom I learned how to approach scientific problems. As one of the earliest researchers in his startup company *Circulomics*, I went through the excitement of conducting research with an objective of product commercialization. My first publication was with Yi, who taught me how to write and prepare data for manuscripts. I am glad to share my six years with Cyrus and Helena, two PhD students from class of 2008 of Biomedical Engineering Department. Research discussions with Tim, Megan, Chris, Vas, Dong Jin, Ye, Liben, Weihua, was very fruitful and scientifically interesting. I wish all the best for those who are to spend their lives for a few years in Wang Lab in the future.

Time to spend with my roommates – Alan, Yun-Ching, Xindong, Lisa, Bryce, Pauline, Brook, Angela, Lin, Karen and Dan – was like oasis in the dessert. I will treasure my memories that I had in our legendary house – "222 UP". I also thank our extended "222 UP" roommates including Tao, Wei-Chiang, Alice, Derrick and many more.

Last but not least, I would thank my family for their warm support during my six years of scientific endeavor. I delicate my dissertation to Ai-Chen Yung, my wife who spent three years with me in Baltimore and is currently working in Taiwan, Qiuping He and Gang Song, my mother and father, as well as my grandparents and relatives in China.

Table of Contents

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Developed in 1983, Polymerase Chain Reaction (PCR) still remains the most widely used technique for genetic biomarker analysis. PCR relies on thermostable polymerase to amplify the target template in an exponential manner, boosting the signal of rare biomarker molecules to detectable levels. Although PCR-based assays are extremely sensitive, they are limited by relatively low specificity and low degrees of multiplexing. Non-specific amplification caused by mispriming is a common problem for PCR-based methods: the polymerase is capable of primer extension in the presence of mismatch hybridization, necessitating further verification processes.(1) Mismatch hybridization poses a limitation to the multiplexing capability of PCR-based methods. The probability of mispriming increases dramatically as more primers are introduced to the reaction. In case of quantitative PCR (qPCR), the current gold standard in PCR-based assays, multiplexing is also limited by the detection method. qPCR, an homogeneous assay that does not rely on downstream separation methods, reports the presence of target through amplified fluorescence signals, which are difficult to separate from one another in the presence of multiple fluorophore species due to their relatively broad spectra.

This dissertation tackles the above problems through two approaches, both of which avoid use of PCR. The first approach is to utilize single

molecule detection (SMD), an extremely sensitive fluorescence detection method that is capable of detecting individual molecules in the sample, to compensate for the reduction in sensitivity in a process without PCR. Instead of PCR, we employed Gap Ligase Chain Reaction (Gap-LCR) to detect the presence of targets in a more specific manner. Due to high specificity of ligase, Gap-LCR can easily distinguish single nucleotide differences in target sequences, making it an ideal assay for genetic mutation detection. The ligation products are further concentrated on quantum dots (QDs) that serve as a nanoconcentrator to further boost fluorescence signals. The Gap-LCR combined with QD-enabled SMD is a highly specific and sensitive assay that distinguishes single nucleotide substitution without a need for PCR. In an attempt to ultimately increase the signal-to-noise ratio of fluorescence reporting, we collaborated with Dr. Mingwei Chen's group, at Tohoku University in Japan, to establish a method to amplify QD fluorescence signal up to 100 times by using nanoporous gold (NPG), a novel nanomaterial that concentrates electromagnetic fields in its pores to enhance the extinction coefficient and the quantum yield of fluorescence molecules. The strong fluorescence enhancement afforded by NPG could potentially reduce the current stringent requirements for SMD optical platforms, making SMD a more ubiquitous fluorescence system in research and clinical laboratories.

In the second approach, we propose an assay, Ligo-miR, which is specifically designed for microRNA, a relatively new class of biomarker. MicroRNAs are small RNA molecules with defined sequences, length, and terminal chemistry, and are estimated to control up to 70% of protein-coding genes through post-transcriptional regulation. Although Ligo-miR is incompatible with other types of biomarkers, such as genetic mutations, methylation, and messenger RNAs, Ligo-miR best meets the requirements of microRNA biomarker detection. Given that most commercially available biomarker detection assays detect a few tens of biomarkers simultaneously, Ligo-miR was also designed to have similar degree of multiplexing – 24 biomarkers simultaneously. Similarly to Gap-LCR, Ligo-miR utilizes the high specificity of ligase to distinguish resembling sequences: microRNAs within the same family (e.g., *let-7a*, *let-7b*, *let-7c*, …). Since microRNAs typically exist at higher concentrations than target DNA sequences (typically 10 - 10,000 microRNAs per cell compared to two copies of, for instance, KRAS sequences), linear amplification is enough to detect targets from tissue samples. Each step of the assay is optimized to provide most convenient, specific, efficient, and unbiased results. The validity of Ligo-miR is also verified using a qPCR assay through differential expression analysis. The Ligo-miR assay developed in my thesis is in the process of being commercialized by *Circulomics Inc*.

This dissertation illustrates that while it may not be feasible to develop an assay that excels in every facet such as sensitivity, specificity, convenience, quantification accuracy and multiplexing, it is possible to develop a practical and informative biomarker detection assay by specifically customizing enzymatic reactions for a certain class of biomarkers. The work in my thesis focuses on biology-driven assay design for biomarker detection, rather than searching for biomarkers after developing a technique that could be used to detect biomolecules.

Chapter 2: Single Quantum Dot-based Ligation Assay for Multiplexed Point Mutation Detection

Song,Y., Zhang,Y. and Wang,T.H. (2013) Single quantum dot analysis enables multiplexed point mutation detection by gap ligase chain reaction. Small, 9, 1096-1105.

2.1 Introduction

DNA single nucleotide substitutions, also known as point mutations, are abnormalities widely found in genomic DNA of various types of cancers. For example, point mutations discovered in *KRAS* genes are closely associated with lung cancer, colorectal cancer, ovarian cancer, and other cancer types.(2) They are therefore valuable genetic markers for cancer diagnostics and prognostics,(3, 4, 5, 6) as well as important predictors of patients' resistance to specific cancer therapies.(7, 8, 9) Currently, the majority of point mutation detection techniques rely on PCR amplification of target sequences from crude genomic DNA samples. Although highly sensitive, PCR based methods are complicated by amplification errors due to mispriming, limited accuracy of discriminating single nucleotide variations, and limited multiplexing capability.(10, 11, 12, 13, 14)

Although a number of alternative PCR-free methods, such as the Invader assay(15) and rolling circle amplification,(16) have been introduced, ligation-based techniques remain the most widely used for point mutation detection

due to their exceptional specificity on base discrimination and robust multiplexing capabilities.(17) A number of variations of ligation assays have been proposed for point mutation detection. Ligase detection reaction (LDR)(18, 19, 20, 21) employs a set of primers to sense the mutation. Only if the primers fully complement the target sequence containing the mutation of interest does the ligase join the two primers together to form ligation products which are then detected using gel electrophoresis or FRET-based approaches.(22, 23, 24) Although it is highly specific in base recognition, LDR has very limited sensitivity. Consequently, LDR is usually combined with PCR that exponentially amplifies the ligation product to a detectable level. The combined PCR-LDR process significantly improves the assay sensitivity but suffers from the complications of PCR. Ligase chain reaction (LCR)(25, 26, 27, 28) has been introduced to enhance the sensitivity of mutation detection by ligation. Instead of using one pair of primers in the case of LDR, LCR uses two pairs of primers to flank both the sense and the antisense strands of DNA targets, generating ligation products that in turn serve as templates for ligation reaction of the next cycle. As a result, the mutation can be easily detected through exponentially amplified ligation products even with gel electrophoresis.(29) Despite high sensitivity, LCR has not been widely adopted for mutation detection. The primers used in LCR would inevitably form primer dimers with blunt ends, which tend to cause false positives due to blunt-end ligation.(30) An improved version of LCR

known as Gap-LCR bypasses the blunt-end ligation by introducing a “gap” between the primers hybridized to the target template.(31, 32) The primers are intentionally designed to form dimers with sticky-ends, thereby eliminating the problem of blunt-end ligation. After filling the gap by DNA polymerase, DNA ligase can seal the nick between primers and generate an allele-specific ligation product. Previous research results suggest that Gap-LCR and allele-specific PCR have similar sensitivity, but Gap-LCR generates less false positives than allele specific PCR when presented with mismatch targets. Gap-LCR achieves this increased specificity by the dual layering of ligase based mismatch discrimination on top of polymerase discrimination.(31) However, most of the ligation-based assays including Gap-LCR rely on cumbersome separation techniques such as gel electrophoresis or solid phase-based purification.(33, 34, 35) Such labor-intensive protocols seriously hinder their applications in routine medical diagnostics procedures.

Although Taqman probes(36, 37, 38, 39) or molecular beacons(40, 41) combined with PCR have enabled separation-free detection of DNA targets in solutions, incomplete quenching of free probes often leads to high fluorescence background and low signal-to-noise ratio.(42) Alternatively, the advancement of single molecule spectroscopy (SMS) and single molecule probe strategies facilitate homogeneous, separation-free detection with high

sensitivity.(43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53) As opposed to conventional ensemble detection methods that measure averaged fluorescence from the entire analyte population, SMS measures fluorescent bursts emitted from individual molecules as they pass through a femtoliter-sized laser detection volume. In SMS, background fluorescence from out-of-focus molecules and scattered light are minimized by a pinhole incorporated to the confocal design. Single molecule coincidence detection(54) is a SMS-enabled approach for sequence-specific detection of single DNA molecules. It employs two differently labeled oligonucleotide probes to search for a specific DNA target. Presence of the target can be determined by coincident fluorescence bursts emitted from the two probes bound to the same target as the probes-target hybrid passes the detection volume of SMS. This strategy permits direct detection of molecular bindings in a solution without the need for separation of free probes from targets. This fluorescence burst coincidence detection method has been successfully applied to detection of specific DNA sequences,(54, 55, 56) DNA methylation,(57) or microRNA expression.(58)

In this report, we introduce a single QD-based multiplexed coincident fluorescence detection method for PCR-free and separation-free detection of point mutations in genomic DNA. The method employs Gap-LCR for multiplexed enzymatic reactions to generate mutation-specific ligation products that are detected by SMS, through single-dot multi-color

fluorescence coincidence analysis, without the need for PCR and molecular separation. In this method, dye-labeled probes are first used for generation of mutation-specific ligation products from genomic DNA templates. Then the ligation products are captured by QDs through biotin-streptavidin binding to form DNA-QD nanocomplexes. Upon illumination, the nanocomplex emits a unique single-dot coincident fluorescence signal that is detected by SMS (**Figure 2.1**A). By performing multiplexed Gap-LCR with multiple sets of differently labeled primers, mutation variants can be determined by specific patterns of coincident signals measured by multi-color SMS (Figure 2.1B). QDs serve as both a high intensity fluorescent label and a nanoconcentrator to capture multiple dye-conjugated ligation products for signal amplification, which facilitates highly accurate detection of coincident fluorescence events. We have demonstrated the method to detect mutations on codon 12 of the *KRAS* gene in human fibroblast and Panc-1 pancreatic cancer cell lines, respectively. The assay has demonstrated exceptional sensitivity, capable of detecting zeptomoles of targets.

2.2. Results

2.2.1. Assay Principles

In the assay, the DNA sample is first subjected to Gap-LCR and then is detected by the coincidence detection platform (Figure 2.1A). In Gap-LCR,

polymerase extends primers only if primers form a perfect match to target templates. Following primer extension, the ligase is introduced to join the primers into ligation products. A set of four primers (two pairs of complementary primers) is used to allow chain reaction. The primer annealing, gap filling and ligation processes are then repeated through thermal cycling in order to exponentially amplify the target. In contrast, the polymerase is not able to elongate the primer to fill the gap if primers and target template form mismatch hybridization. Consequently, no ligation product is formed in the case of mismatch. Since one of the primers is labeled with biotin and another labeled with fluorophore, ligation products which are generated by Gap-LCR are labeled with both a biotin and a fluorophore. Such ligation products are captured by streptavidin-functionalized QDs to form DNA-QD nanocomplexes that are subsequently analyzed using coincidence detection. In the coincidence detection platform, the laser beam is focused into a femtoliter-sized volume from where fluorescence signals are collected. When a DNA-QD nanocomplex diffuses into the laser focal volume of the SMS platform, both QD and the fluorophore labeled on the ligation products are excited, emitting a coincident fluorescence burst that indicates the presence of target template of interest. A pinhole placed in front of the detectors is used to reject fluorescence emitted by out-of-focus molecules, allowing high signal-to-noise ratio detection and preventing detection of false-positives (**Figure S2.1**).

In order for multiplexed detection of single nucleotide polymorphisms, a set of five primers that include three common primers and two variant-specific primers (**Table 2.1**) are used to simultaneously detect mutant and wild type target (Figure 2.1B). Mutant-specific and wild type-specific primers are labeled with different species of fluorophores. Depending on the genotype of target template (homozygous mutant, wild type or heterozygous mutant), DNA-QD nanocomplexes of different color-combinations are generated and are detected using the coincidence detection platform.

2.2.2. Assay Validation

We evaluated our assay by detecting *KRAS* codon 12 GGT → GAT mutation using the sample containing mutant targets. Four primers including mutant-specific primer Apr and common primers Cpr1-3 were used. Apr and Cpr1 were labeled with Alexa488 and biotin, respectively (Table 2.1). Ligation products were generated by Gap-LCR and were captured by QDs. As a result, coincident fluorescence bursts were observed in both QD and Alexa488 optical channels (marked by stars), indicating the presence of DNA-QD nanocomplexes (**Figure 2.2**A, B). In the case of mismatch control, where the sample contained only the wild type *KRAS* target, no ligation product was generated as evidenced by the absence of coincident fluorescence burst events (Figure 2.2C, D). As mentioned earlier, QDs resulted in stronger fluorescence bursts due to their high extinction coefficient and quantum yield. In the

meantime, the fluorescence burst size of Alexa488 was significantly increased in case of perfect match because QDs brought multiple Alexa488 dyes together to pass through the detection volume at the same time. Therefore, QDs did not only serve as nanosensors that confirmed the presence of ligation products through coincidence detection, but also act as target concentrators that greatly enhanced the signals from labeled fluorophores. The number of coincident events was plotted as a function of the threshold applied to Alexa488 signal (Figure 2.2E). Since the primers are specific to the *KRAS* mutant, only the mutant samples exhibited detectable coincident events. The wild type in this case displayed negligible coincident signals. The specificity of the wild type primers was also tested, in which case the coincident events were observed in the wild type but not in the mutant samples (**Figure S2.2**).

In order to further verify the results obtained using coincidence detection, the cross-correlation function(59) $G(\tau)$ was calculated based on SMS data of Alexa488 and QD according to

$$G(\tau) = \frac{\left\langle \delta I_{A488}(t)\delta I_{QD}(t+\tau)\right\rangle}{\left\langle I_{A488}(t)\right\rangle\left\langle I_{QD}(t)\right\rangle} \qquad (1)$$

where τ was the lag time, I_{A488} and I_{QD} were the fluorescent intensities from the Alexa488 channel and the QD channel respectively, and δI_{A488}, δI_{QD} were the deviation of Alexa488 and QD655 intensities from their respective means. As shown in Figure 2.2F, only the mutant sample exhibited a strong

correlation between the signals from Alexa488 and QD as evidenced by the pronounced peak, which suggested that fluorescence bursts were simultaneously detected from both channels as a result of the ligation products being generated and captured by the QDs. In contrast, the wild type showed minimum cross-correlation because the Alexa488 and QDs moved independently through the detection volume in the absence of ligation products.

2.2.3. Assessment of Analytical Specificity

Figure 2.3 shows the intensity histograms of Alexa488 fluorescence bursts detected from three different samples including mutant, wild type and no-template control. Since ligation was prohibited in the no-template control sample, the detected fluorescence bursts should originate only from the free Alexa488-labeled primers. The free primers exhibited an intensity level of lower than 140 photon counts/ms and a similar result was observed in the wild type experiment. In contrast, fluorescence bursts with a much enhanced intensity of 500 photon counts/ms or higher were detected in the mutant sample. The percentages of Alexa488 fluorescence bursts that were coincident with QD fluorescence bursts were calculated as also shown in Figure 2.3. A small number of statistical fluorescence coincident events (2-6%) were counted in both the wild type and the control experiments when using a small intensity threshold. Since QDs serve as a nanoconcentrator to facilitate

signal enhancement, a higher threshold can be imposed to achieve effective differentiation of binding-induced coincident signals and the statistical coincident signals, greatly enhancing the specificity and accuracy in mutation detection. In the current experiment, the statistical background coincident signals, which may lead to false positives, can be completely removed when a higher threshold (≥ 120 photon counts/ms) is used, thereby achieving near 100% specificity. The same conclusion can be drawn for detection of wild type target as shown in Alexa546 intensity histogram (Figure S2.3).

2.2.4. Assessment of Analytical Sensitivity

In order to access the sensitivity of the assay, we tested the assay with various amounts of target input (**Figure 2.4**). A dilution series from 17 attomole to 170 zeptomole of mutant and the wild type were used as target templates. Instead of the Alexa488-labeled, mutant specific Apr primer, we used wild type specific primer Gpr, which was labeled with Alexa546, to detect the wild type target. As expected, samples with larger amount of target input showed higher coincidence count at any threshold due to larger amount of ligation products generated in the Gap-LCR reaction. Signals from mutant as well as wild type samples with zeptomoles of target input were routinely detected above the background of no-template control, demonstrating excellent sensitivity.

The proposed QD-SMS Gap-LCR assay has been applied for *KRAS* genotyping in human foreskin fibroblast and pancreatic cancer cell lines (Panc-1). Human foreskin fibroblast carries wild type GGT at *KRAS* codon 12 while Panc-1 cancer cells carry GGT → GAT mis-sense mutation. DNA was first extracted from the cell lines. Then, the mutant was amplified with mutant-specific primers and the wild type was amplified with wild type-specific primers by Gap-LCR. Compared to no-template control, high coincident counts were observed in positive samples over a wide range of threshold, indicating successful identification of particular genotypes (**Figure 2.5**).

2.2.5. Multiplexed *KRAS* Mutation Detection

In the QD nano-assay, the high detection specificity and sensitivity complements its multiplexing capability. The assay was able to differentiate between the homogeneous mutant, the heterozygous mutant and the wild type by analyzing the sample using differently labeled mutant specific primers and wild type-specific primers in a single reaction (Figure 2.1B). The experiment was performed at a ligation temperature of 65.1 °C, optimized for multiplexed ligation reaction for *KRAS* codon 12 (Supporting Information and **Figure S2.4**). Four samples, including homozygous mutant, heterozygous mutant, wild type, and no-template control, were analyzed and coincidence detection was conducted with SMS via two dual-color detection

settings (see Materials and Methods). For the homozygous mutant, the coincident events were observed between QD and Alexa488, but not between QD and Alexa546 (**Figure 2.6**A and B). In contrast, the coincident events were only observed between QD and Alexa546 but not between QD and Alexa488 in the case of wild type sample (Figure 2.6C and D). For heterozygous mutant sample, coincident events were observed with both combinations due to the presence of both mutant and wild type alleles (Figure 2.6E and F). In **Figure 2.7**, we also statistically verified that coincidence fluorescence bursts obtained from samples containing perfect match target templates were significantly higher than samples which didn't. As a result, the combination of the coincidence analysis provided unambiguous information for determining the mutational status of the sample under investigation. This demonstrates that the assay is capable of detecting point mutations in a multiplexed manner with high specificity.

2.3. Discussion

Herein, we demonstrated a highly sensitive and specific QD nano-assay which reveals genotype of the target biomarker in a multiplexed manner. Our single molecule approach provides significant values over the conventional ensemble measurements. Firstly, perfect match samples with Gap-LCR reaction (leading to labeled ligation product) and those without (leaving

behind free labeled primers) result in same ensemble fluorescence intensities. Therefore, conventional ensemble measurements need to rely on an additional separation step such as gel electrophoresis or bead-based purification in order to isolate ligation products from free primers to distinguish signal resulting from ligation products from the background. Alternatively, SMS can be used for the examination of individual molecules in a homogeneous environment. Therefore, ligation products and free primers labeled with the same fluorophore could be easily distinguished from one another using coincidence detection in a separation-free manner, which greatly improves the sensitivity and simplifies the procedures of the assay.

Secondly, multiplexing experiment using multiple species of fluorophores also greatly benefit from a SMS-enabled approach compared to conventional ensemble measurement. While even a minor degree of fluorophores' emission spectra bleed through from one detection channel to the other might result in false positive detection in ensemble measurements, SMS is capable of circumventing the problem. Conventional fluorescence measurement relies on analog fluorescence acquisition. Therefore, when multiple species of fluorophores are present in the sample, it is difficult decouple the signal obtained from targeted fluorophores and bleed through from other fluorophore species. For instance, we confirmed that there is <5% spectra bleed through of Alexa488 fluorescence signal into Alexa546 channel in

ensemble measurement (data not shown). On the other hand, SMS decouples the analog ensemble fluorescence signal into digital signals from individual molecules. Therefore, by setting the threshold, it becomes possible to totally eliminate weak bleed through signals without losing any targeted fluorescence bursts, which are significantly stronger than bleed through signals and background noise (Figure 2.2B).

Multiplexing capability of separation-free assays, or the number of genetic markers which can be investigated in single reaction, is dependent on availability of fluorophore combinations which can be cleanly separated by optical filters. On one hand, since most organic fluorophores have relatively broad emission spectra, it becomes challenging to optically separate more than a few fluorophores at the same time. On the other hand, since the emission spectra of QDs are narrow and can be easily separated from each other, QD biosensor could be a powerful solution to boost the multiplexing capability of separation-free assays. Further improvement in conjugation techniques would prove beneficial in order to completely substitute chemical fluorophores labeled on DNA with QDs. Specifically, more variations of suitable conjugation method are required so that each species of QD can be easily, stably, and specifically labeled to target DNA molecules without affecting the biochemical reaction and components in the solution.

2.4. Conclusions

In summary, we demonstrated single-QD, multiple-color coincident fluorescence analysis that facilitates multiplexed and separation-free detection of point mutations. By combining the high specificity of Gap-LCR and high sensitivity of QD-SMS, we successfully detected zeptomoles of *KRAS* codon 12 mutation variants with close to 100% specificity. Such high sensitivity eliminated the requirement for PCR pre-amplification from genomic DNA. In addition, a tedious separation step commonly required by ligation-based mutation assays was circumvented by the integration of coincidence detection, which identifies specific target sequence by detecting coincident multi-color fluorescence bursts in homogeneous environment. Such a highly sensitive, specific, multiplexed separation-free genotyping method would prove valuable as novel single molecule approach for genetic mutation detection.

2.5. Experimental Section

DNA Extraction. Genomic DNA was extracted from human foreskin fibroblast (CRL-2522, ATCC, Manassas, VA) and Panc-1 cancer cell lines using Biosprint 15 DNA extraction kit (Qiagen, Inc., Venlo, Netherlands). The extracted DNA was stored at 4 °C until use.

Oligonucleotides. All the synthetic oligonucleotides were purchased from the Integrated DNA Technology and the sequences were listed in Table 2.1. Apr and Gpr were labeled with Alexa488 and Alexa546, respectively, at their 5' terminals. Cpr1 and Cpr3 were labeled with phosphate at their 5' terminals for ligation reaction to take place. Primer Cpr1 was labeled with biotin at its 3' terminal for coupling with streptavidin-coated QD. All primers were HPLC purified.

Gap-LCR Reaction. Gap-LCR was conducted in 25 μl of reaction volume containing ThermoPol Buffer (1X); *Taq* Ligase Buffer (0.25X); NAD β (1mM); dGTP (10 μM); dATP (10 μM); $MgCl_2$ (3 mM); *Taq* DNA Polymerase (1.25 units, New England Biolabs); *Taq* DNA Ligase (10 units, New England Biolabs); various amounts of mutant and/or wild type templates (synthetic or genetic DNA extracted from cell lines); common primers (60 nM, Cpr1 – 3) as well as discrimination primers (Apr and/or Gpr). The mixture was first incubated at 94 °C for 2 minutes, and then thermal cycling was conducted for 60 cycles at 94 °C for 30 seconds (denaturation step) and 65.1 °C for 60 seconds (ligation step) unless otherwise noted (Supporting Information). After thermal cycling, *Taq* DNA polymerase and *Taq* DNA ligase were inactivated by incubating the sample at 94 °C for 30 minutes. Ligation products were stored at 4 °C until use.

Coupling of QD and Ligation Products. Qdot 655 streptavidin conjugate (QD655), which has a peak emission wavelength at 655 nm, was purchased from Invitrogen (Carlsbad, CA). QD was first diluted to 1 nM with phosphate-buffered saline without calcium & magnesium (1X, Invitrogen). The ligation products were then mixed with QD at 30:1 biotin:QD ratio in order to generate DNA-QD nanocomplex. Since each QD had 40-80 biotin binding sites, all biotin-labeled DNA molecules could be captured by QDs (supporting information). After 20 minutes incubation in dark at room temperature, the nanocomplex was further diluted to 50 pM using PBS. Later, the samples were dispensed into 96-well plate with glass bottom for SMS measurements. All tubes were washed with BSA solution (0.1 %) and were baked at 70°C degrees for 15 min before use.

Coincident Fluorescence Detection. Coincident fluorescence detection was conducted using Zeiss LSM 510 Confocal microscope (Figure S2.1). 488 nm Argon laser and 543 nm Helium-Neon Laser were used to excite Alexa488 and Alexa546, respectively. Both lasers were capable of exciting QD655. The 488 nm laser and 543 nm laser were kept at 7.2 mW and 1.2 mW, respectively, at all time. A C-Apochromat 40X/1.2 water immersion objective was used to focus the laser into the sample contained in the 96-well plate and to collect emitted fluorescence. A 90 □m pinhole (PH) was placed after dichroic mirror (DM) to reject background fluorescence from out-of-focus

molecules and scattered light. Fluorescence emitted from in-focus molecules was separated by dichroic mirrors to fluorescence emitted from fluorophore labeled on ligation product (Alexa488 and/or Alexa546) and from QD. Fluorescence light was then respectively filtered by emission filters (EF) and subsequently detected by avalanche photodiode (APD). Measurements were conducted for 180 seconds for each sample.

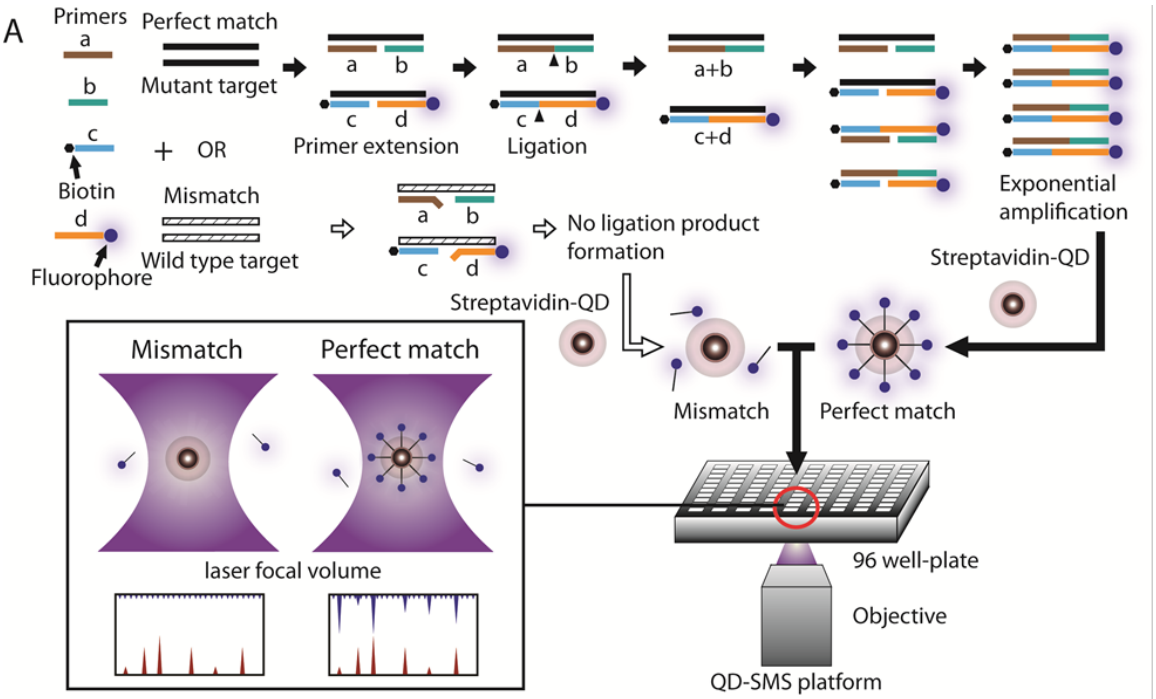
Coincidence Analysis. Acquired data were analyzed by a customized program written in LabView (National Instrument, Autstin, TX). Individual fluorescence bursts were first identified by applying threshold to raw single molecule data. Fluorescence signals which exceeded the threshold were recognized as true fluorescence bursts. Coincident fluorescence bursts were then identified as a pair of simultaneous fluorescence bursts from both QD and Alexa488/546 channels.

2.6. Tables and Figures

Table 2.1. Oligonucleotide.

Code	Description	Sequence (5' to 3')
KAS	A segment in mutant type *KRAS* codon 12 gene sense strand (GGT → GAT)	AAT ATA AAC TTG TGG TAG TTG GAG CTG ATG GCG TAG GCA AGA GTG CCT
KAA	A segment in mutant type *KRAS* codon 12 gene anti-sense strand (Complementary to KAS)	AGG CAC TCT TGC CTA CGC CAT CAG CTC CAA CTA CCA CAA GTT TAT ATT
KGS	A segment in wild type *KRAS* codon 12 gene sense strand	AAT ATA AAC TTG TGG TAG TTG GAG CTG GTG GCG TAG GCA AGA GTG CCT
KGA	A segment in wild type *KRAS* codon 12 gene anti-sense strand (Complementary to KGS)	AGG CAC TCT TGC CTA CGC CAC CAG CTC CAA CTA CCA CAA GTT TAT ATT
Apr	Alexa488 labeled discrimination primer for GGT → GAT mutant *KRAS* detection	A488-CAC TCT TGC CTA CGC CAT C
Gpr	Alexa546 labeled discrimination primer for wild type *KRAS* detection	A546- CAC TCT TGC CTA CGC CAC C
Cpr1	Common primer 1 labeled with phosphate and biotin	p CTC CAA CTA CCA CAA GTT TAT ATT-Biotin
Cpr2	Common primer 2	AAT ATA AAC TTG TGG TAG TTG GAG CT
Cpr3	Common primer 3 labeled with phosphate	p TGG CGT AGG CAA GAG TG

Figure 2.1. Schematic drawing of the point mutation detection assay. (A) The mutant-specific primer is labeled with a fluorophore and a common primer is labeled with biotin. In case of a perfect match (flow illustrated by black arrows), primers hybridize to the mutant target template (black strand) with a gap in-between. The polymerase fills the gap and then the nick (black triangle) is sealed by the ligase. The ligation product serves as a target template in the next cycle of the thermal cycling, leading to exponential increase of target template. After Gap-LCR reaction, streptavidin-coated QDs are introduced to form DNA-QD nanocomples, which give rise to coincident fluorescence bursts when analyzed in single molecule coincidence detection platform. In case of a mismatch (flow illustrated by white arrows), the primers do not hybridize stably to wild type target templates (striped strand). As a result, no ligation product is formed and captured by QD, as evidenced by absence of coincident fluorescence bursts. (B) Multiplexed detection of homozygous mutant, wild type and heterozygous mutant. Mutant and wild type-specific primers are introduced simultaneously. The presence of mutant is evidenced as coincident fluorescence burst from Alexa488 (blue) and QD (red) channel. Similarly, presence of wild type target is evidenced as coincident fluorescence burst from Alexa546 (green) and QD (red) channel.


A
Primers
a
b
c
Biotin
d
Fluorophore
Perfect match
Mutant target
+ OR
Mismatch
Wild type target
Primer extension
Ligation
a+b
c+d
Exponential amplification
No ligation product formation
Streptavidin-QD
Streptavidin-QD
Mismatch
Perfect match
Mismatch
Perfect match
laser focal volume
96 well-plate
Objective
QD-SMS platform

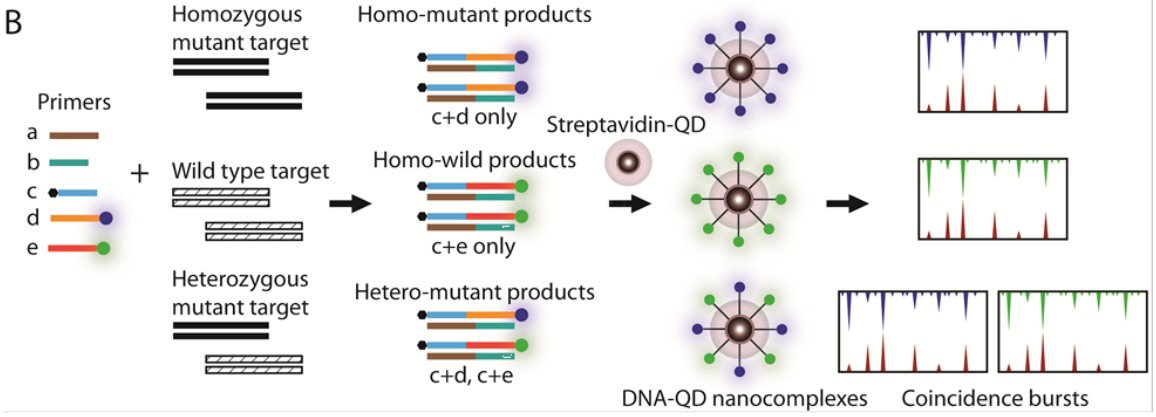

B
Primers
a
b
c
d
e
Homozygous mutant target
Wild type target
Heterozygous mutant target
Homo-mutant products
c+d only
Homo-wild products
c+e only
Hetero-mutant products
c+d, c+e
Streptavidin-QD
DNA-QD nanocomplexes
Coincidence bursts

Figure 2.2. Detection of *KRAS* mutant target.

(A) Mutant was detected as evidenced by coincident fluorescence bursts (marked with stars) from Alexa488 (blue) and QD (red) channels. (B) Enlarged view of coincident fluorescence bursts from panel (A). Regions specified with green slits were enlarged. Because each QD captures multiple ligation products, signal from Alexa488 was significantly enhanced (arrows). (C) Mutant-specific primers do not form perfect match with wild type target. Therefore, no coincident fluorescence bursts were observed. (D) Enlarged view of QD fluorescence bursts from panel (C). Arrows indicate unamplified Alexa Fluor488 fluorescence bursts observed from free primers. (E) Coincident fluorescence bursts at various thresholds. (F) Cross-correlation analysis of QD and Alexa488 signals reveals a strong correlation for the mutant but not wild type samples.

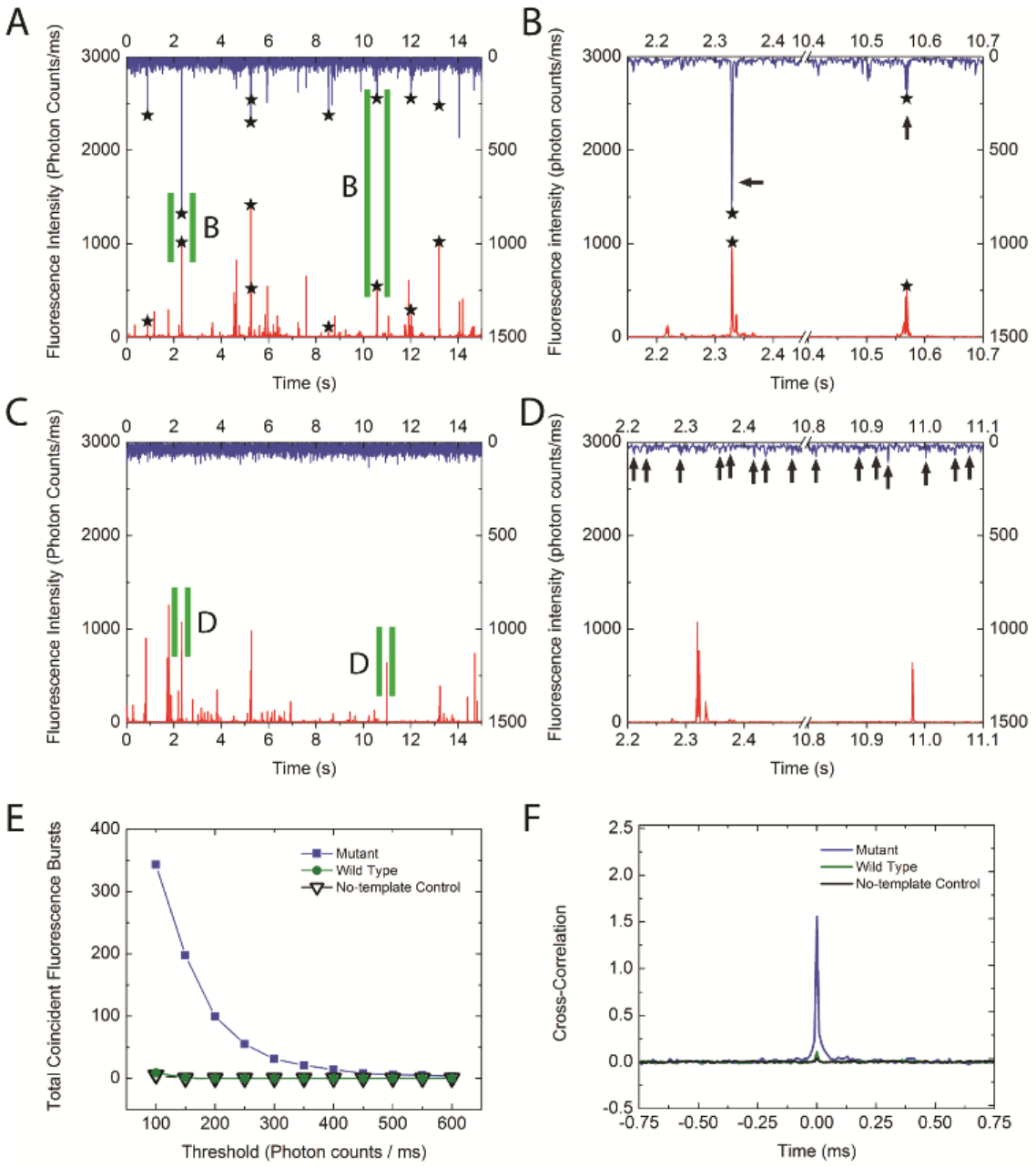

A
B
C
D
E
F
Fluorescence Intensity (Photon Counts/ms)
Fluorescence intensity (photon counts/ms)
Time (s)
B
D
Total Coincident Fluorescence Bursts
Threshold (Photon counts / ms)
Mutant
Wild Type
No-template Control
Cross-Correlation
Time (ms)

Figure 2.3. Specifcity assessment of *KRAS* wild type detection.

Histograms of Alexa488 fluorescence bursts detected from perfect match (mutant sample, shown in densely striped bars), mismatch (wild type sample, shown in sparsely striped bars) and no-template control samples (black bars) in 180 second-measurements. The number above each bar presents the percentage of Alexa488 fluorescence bursts that were coincident with QD fluorescence bursts.

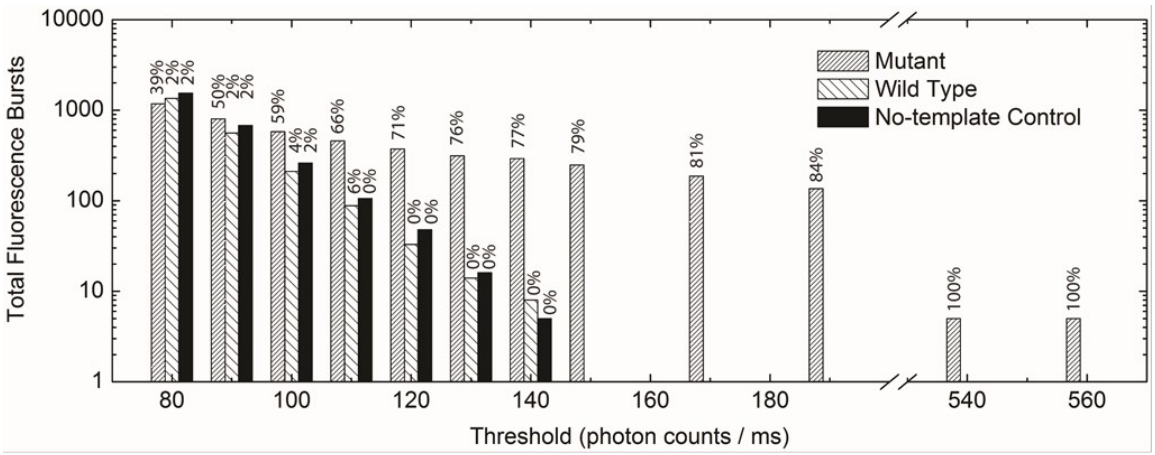

Total Fluorescence Bursts
10000
1000
100
10
1
Mutant
Wild Type
No-template Control
39%
2%
2%
50%
2%
2%
59%
4%
2%
66%
6%
0%
71%
0%
0%
76%
0%
0%
77%
0%
0%
79%
81%
84%
100%
100%
80
100
120
140
160
180
540
560
Threshold (photon counts / ms)

Figure 2.4. Sensitivity assessment.

Detection of serial dilutions of (A) mutant target and (B) wild type target. The normalized coincidence count is defined as the number of coincidence bursts divided by the number of total QD bursts observed within the respective detection period.

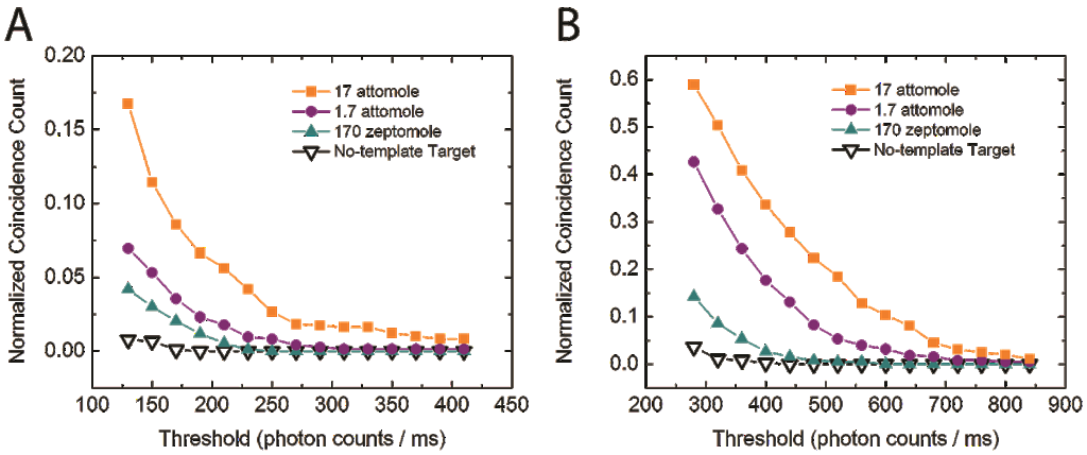

A
Normalized Coincidence Count
0.20
0.15
0.10
0.05
0.00
100
150
200
250
300
350
400
450
Threshold (photon counts / ms)
17 attomole
1.7 attomole
170 zeptomole
No-template Target
B
Normalized Coincidence Count
0.6
0.5
0.4
0.3
0.2
0.1
0.0
200
300
400
500
600
700
800
900
Threshold (photon counts / ms)
17 attomole
1.7 attomole
170 zeptomole
No-template Target

Figure 2.5. Detection of *KRAS* sequence from cell line DNAs.

KRAS codon 12 point mutation detection from genomic DNA samples. (A) GGT → GAT mutant detection from genomic DNA extracted from pancreatic cancer cell line (Panc-1). (B) Wild type target detection from genomic DNA extracted from human foreskin fibroblast.

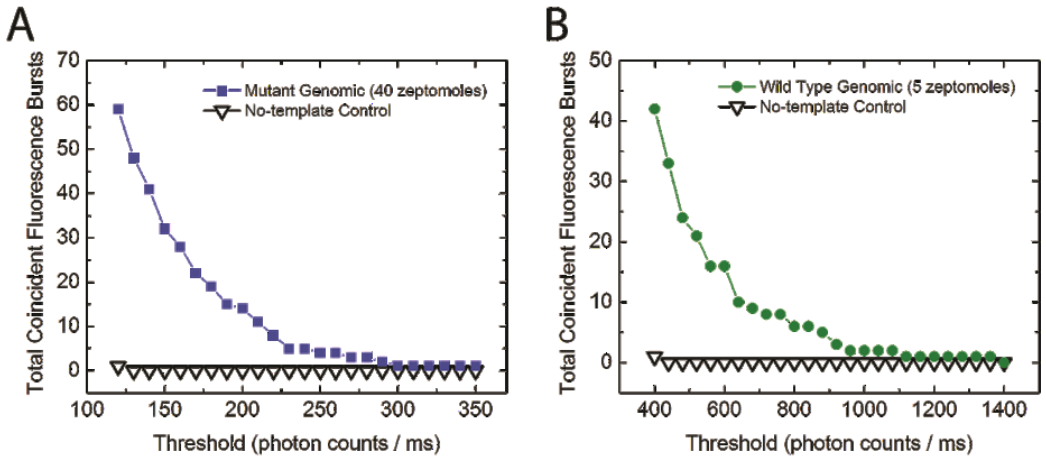

A
Total Coincident Fluorescence Bursts
Mutant Genomic (40 zeptomoles)
No-template Control
Threshold (photon counts / ms)
B
Total Coincident Fluorescence Bursts
Wild Type Genomic (5 zeptomoles)
No-template Control
Threshold (photon counts / ms)

Figure 2.6. Multiplexed point mutation detection.

Coincident fluorescence bursts (marked with starts) from Alexa488 (blue) and QD (red) channels suggest the successful detection of mutant target in homozygous mutant sample (A) and heterozygous mutant sample (E). Coincident fluorescence bursts from Alexa546 (green) and QD (red) channels suggested the successful detection of wild type target in wild type sample (D) and heterozygous mutant sample (F). Coincident Fluorescence bursts were not detected in absence of specific targets (B and C).

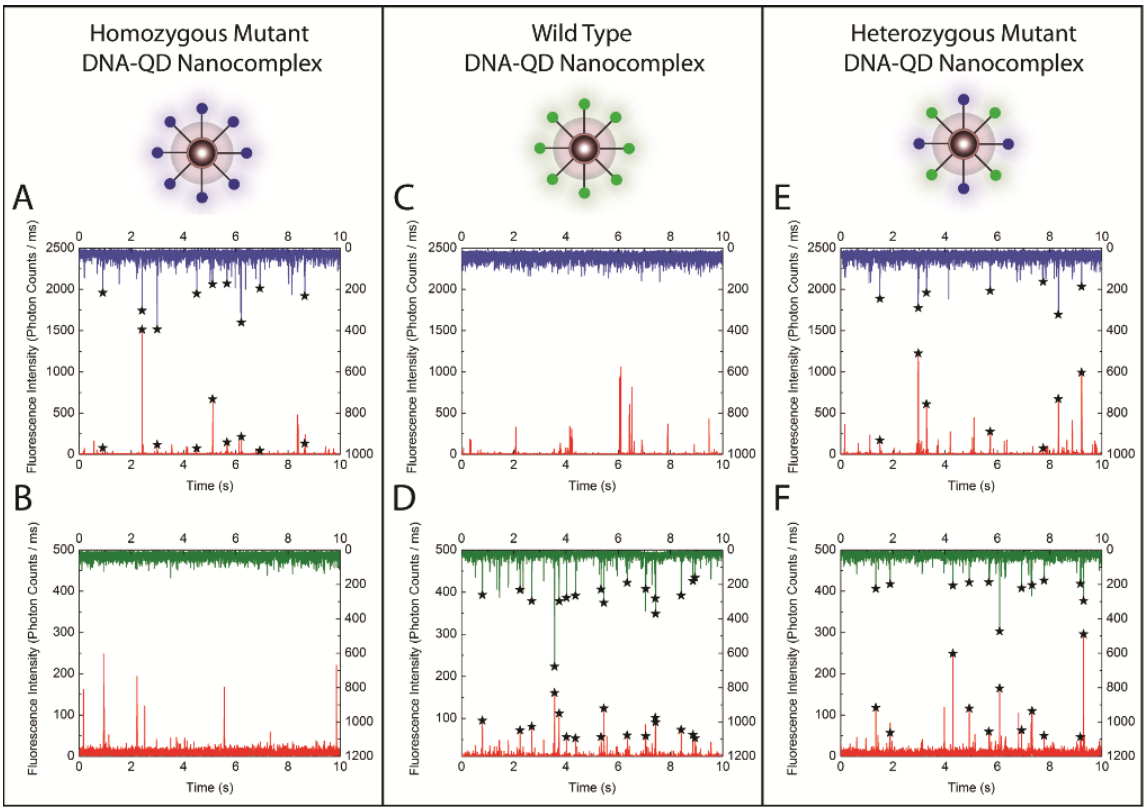
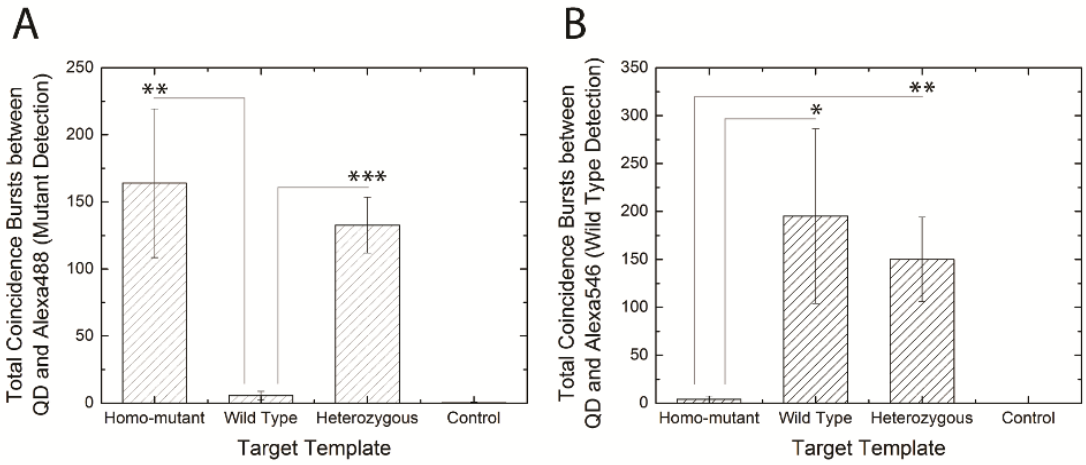

Homozygous Mutant
DNA-QD Nanocomplex
Wild Type
DNA-QD Nanocomplex
Heterozygous Mutant
DNA-QD Nanocomplex
A
B
C
D
E
F
Fluorescence Intensity (Photon Counts / ms)
Time (s)

Figure 2.7. Analysis of multiplexed point mutation detection.

The asterisk, double asterisk, and triple asterisk indicate that p-value is less than 0.05, 0.01, and 0.001, respectively, in Student's one-tailed T-test. The error bar indicates standard deviation of three measurements. (A) Detection of mutant target. Only homozygous mutant sample and heterozygous mutant sample showed significant number of coincident fluorescence bursts. (B) Detection of wild type target. Only wild type sample and heterozygous mutant sample showed significant number of coincident fluorescence bursts.


A
Total Coincidence Bursts between QD and Alexa488 (Mutant Detection)
250
200
150
100
50
0
**
***
Homo-mutant
Wild Type
Heterozygous
Control
Target Template
B
Total Coincidence Bursts between QD and Alexa546 (Wild Type Detection)
350
300
250
200
150
100
50
0
*
**
Homo-mutant
Wild Type
Heterozygous
Control
Target Template

2.7. Supplementary Information

Ligation temperature optimization. In order to optimize the ligation temperature in the assay, we conducted multiplexed point mutation detection using three ligation temperatures, 64.4, 65.1, and 65.8 °C (Figure S2.4). The following four samples were tested with the assay: homozygous mutant, wild type, heterozygous mutant, and no-template control. At 64.4 °C, the assay detected mutant sequence from homozygous and heterozygous mutant samples and wild type sequence from wild type and heterozygous mutant samples. However, template-independent amplification was observed from no-template control sample at this temperature (Figure S2.4A). Template-independent amplification was suppressed when the ligation temperature was increased to 65.1 °C or higher (Figure S2.4A, B). When the ligation step was conducted at 65.8 °C, much less coincidence bursts were observed from all four samples because of the reduced activity of *Taq* DNA Ligase and weaker primer-target template hybridization. Therefore, we concluded that 65.1 °C was the optimum ligation temperature for multiplexed detection of mutant and wild type *KRAS* sequences.

DNA-QD conjugation. It is important that all biotin-labeled DNA molecules are captured by QDs so as to prevent false-negative reaction results. In order to accomplish this objective, we incubated DNA with QD whose total biotin-binding sites are significantly more than all biotin-labled DNA molecules in

solution. Since each QD has 40-80 biotin binding sites, we incubated biotin-labeled DNA and QD at 30:1 ratio. Under this condition, we expect each QD molecule to capture approximately 30 DNA molecules in fairly consistent manner. Inconsistent DNA-QD conjugation could possibly result from either (1) inconsistent QD sizes (and hence surface area) or (2) insufficient mixing. However, neither of them is likely. (1) We expect the QD size to be fairly uniform because QD emission spectra, which are dependent on QD sizes, are narrow. (2) Mixing is expected to be efficient and uniform because both DNA and QD are highly water soluble and are evenly distributed in the solution before mixing.

Figure S2.1. Multi-color fluorescence burst coincidence analysis platform.

The laser beams are focused into the sample solution contained in 96 well-plate by an objective. Fluorescence emitted from single molecules is collected by the objective, separated by dichroic mirror (DM) by spectrum, filtered by emission filter (EF) and detected by avalanche photodiode (APD). A Pinhole (PH) is used to reject background fluorescence from out-of-focus molecules.

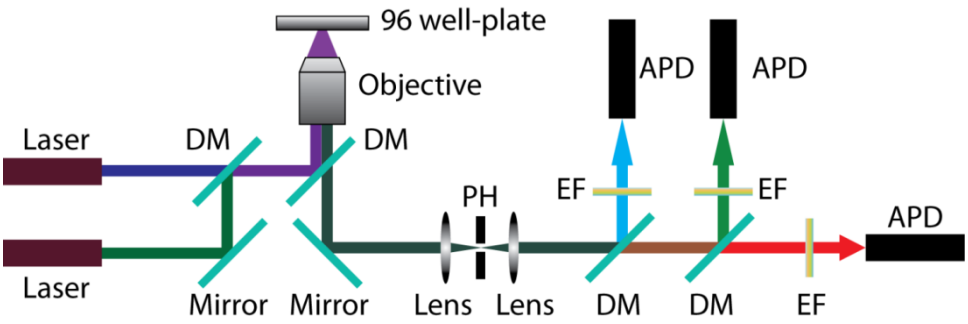

96 well-plate
Objective
APD
APD
Laser
DM
DM
PH
EF
EF
APD
Laser
Mirror
Mirror
Lens
Lens
DM
DM
EF

Figure S2.2. Detection of wild type *KRAS* sequence

Wild type-specific primer set (Cpr1-3 and Gpr) was used. Only wild type sample showed significant coincident fluorescence bursts.

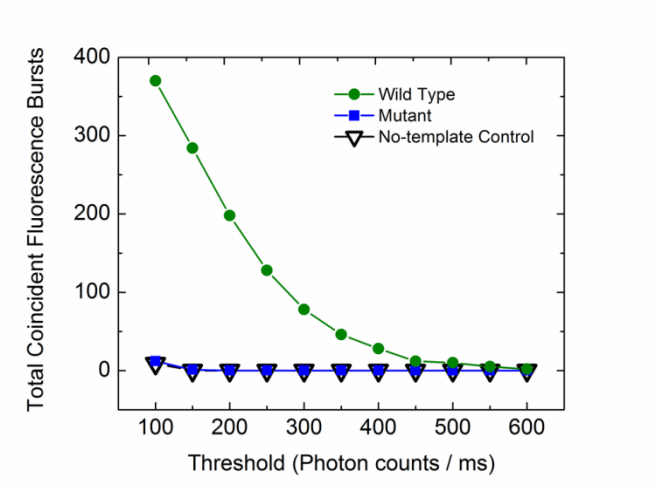

Total Coincident Fluorescence Bursts
400
300
200
100
0
100
200
300
400
500
600
Threshold (Photon counts / ms)
Wild Type
Mutant
No-template Control

Figure S2.3. Specifcity assessment of *KRAS* wild type detection.

Histograms of Alexa546 fluorescence bursts detected from perfect match (wild type sample, shown in sparsely striped bars), mismatch (mutant sample, shown in densely striped bars) and no-template control samples (black bars) in 180 second-measurements. The number above each bar presents the percentage of Alexa546 fluorescence bursts that were coincident with QD fluorescence bursts.

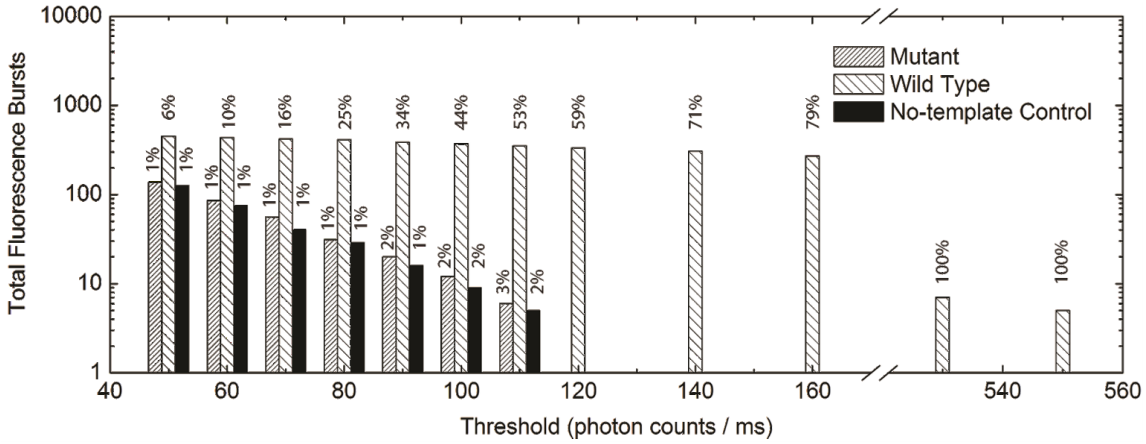
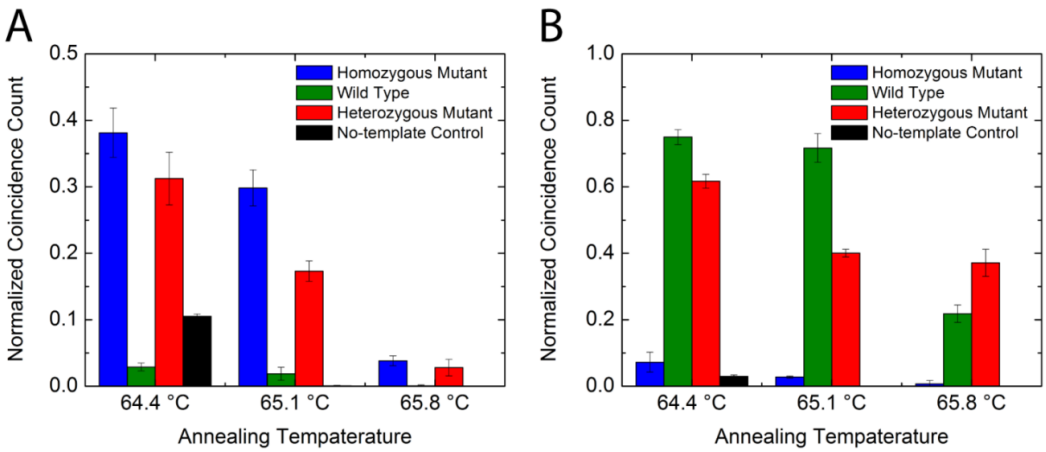

Total Fluorescence Bursts
10000
1000
100
10
1
40
60
80
100
120
140
160
540
560
Threshold (photon counts / ms)
Mutant
Wild Type
No-template Control
1%
6%
1%
1%
10%
1%
1%
16%
1%
1%
25%
1%
2%
34%
1%
2%
44%
2%
3%
53%
2%
59%
71%
79%
100%
100%

Figure S2.4. Ligation temperature optimization for multiplexed point mutation detection.

Detection of mutant target (A) and wild type target (B) using three ligation temperatures. The error bar indicates standard deviation of three measurements.


A
Normalized Coincidence Count
0.5
0.4
0.3
0.2
0.1
0.0
Homozygous Mutant
Wild Type
Heterozygous Mutant
No-template Control
64.4 °C
65.1 °C
65.8 °C
Annealing Tempaterature
B
Normalized Coincidence Count
1.0
0.8
0.6
0.4
0.2
0.0
Homozygous Mutant
Wild Type
Heterozygous Mutant
No-template Control
64.4 °C
65.1 °C
65.8 °C
Annealing Tempaterature

Chapter 3: Enhancement of Quantum Dot Fluorescence by Nanoporous Gold (NPG)

Zhang,L., Song,Y., Fujita,T., Zhang,Y., Chen,M. and Wang,T.H. (2014) Large enhancement of quantum dot fluorescence by highly scalable nanoporous gold. Adv Mater, 26, 1289-1294.

3.1 Introduction

Nanoengineered metallic materials have been shown to have a number of exclusive physicochemical properties not available at neither larger (micro- and macroscopic) nor smaller (molecular) scales. Recently, these materials in particular have drawn significant attention due to their capability to enhance fluorescent signals of nearby fluorescent species through a phenomenon known as metal enhanced fluorescence (MEF). MEF originates from the localized surface plasmon resonance (LSPR), a collective oscillation of conduction-band electrons that can modify both the extinction coefficient and quantum yield of adjacent fluorescent molecules/species.(60) The extinction coefficient is a function of the electromagnetic field intensity experienced by the fluorescent molecules, and under an enhanced electromagnetic field, fluorescent molecules absorb photons and promote electrons into excited states at accelerated rates. LSPR of metallic nanostructures can also modify the quantum yield of fluorescent species by increasing their radioactive decay

rate. The combined enhancement of extinction coefficient and quantum yield can result in significantly strengthened fluorescence from fluorescent species.

Most MEF materials developed thus far are targeted for enhancing fluorescence from organic fluorophores,(61, 62, 63, 64, 65) and few have been developed for inorganic fluorescent species such as quantum dots (QDs). (66, 67, 68, 69, 70, 71) QDs are distinguished by size-tuneable spectra, large Stokes shifts, high quantum yields and great photostability. (72) These properties make QDs a desirable candidate for applications in biomolecular sensing(73, 74) and cellular imaging.(75, 76) Although MEF can improve the detection sensitivity and imaging quality, wide implementation in bioassays is obstructed by technological limitations. In particular, fabrication of metallic nanostructures for MEF is often not scalable due to the requirement of sophisticated equipment and highly specialized personnel.(69) Furthermore, conventional MEF materials, such as nanoantennas,(64, 77) nanorods(67) and nanoparticles(63, 65, 78, 79, 80) induce large fluorescence enhancements only at nanoscopic “hot spots”. (78) Since the sparse hot spots constitute only a small fraction of the surface area, the majority of molecules reside outside the active region and do not experience MEF. This results in a low overall enhancement that is impractical for ensemble fluorescence assays, such as microarrays and immunoassays, in which signals are measured across the entire substrate area.

In this report, we demonstrate that a three-dimensional free-standing nanoporous gold (NPG) film can greatly enhance QD fluorescence. NPG was fabricated by a one-step etching process, eliminating the need for a sophisticated nanolithography facility. Strong near-field excitation induced by NPG enabled ~100-fold enhancement of QD fluorescence, the highest demonstrated value to our knowledge. Strong MEF was observed across the entire NPG substrate comprising dense and uniform nanoporous structures. The broad active region of MEF makes NPG perfectly appropriate for ensemble fluorescence assays. In addition, MEF by NPG significantly improved fluorescence imaging of QDs at the single dot level. Different QDs were found to have different optimum pore sizes for maximum enhancement; the QD525 was best enhanced by 38 nm NPG whereas QD605 by 51 nm NPG. Importantly, the pore size of NPG can be tailored to a specific QD species by simply adjusting etching time.

3.2. Results and Discussion

We fabricated NPG films with various pore sizes ranging from 17 nm to 71 nm in average diameter. The NPG films were prepared by selectively etching Ag from 700 nm thick Ag65Au35 (at. %) sheets using saturated nitric acid at room temperature (**Figure 3.1a and b**).(81) As nitric acid etches silver atoms away from the alloy, a bicontinuous nanoporous structure with interconnected nanometer-sized pores and gold ligaments is formed by self-

assembly of residual gold atoms (**Figure 3.1c**). The as-prepared NPG films were carefully rinsed with distilled water. The pore sizes, defined by the equivalent diameters of nanopore channels or gold ligaments, can be tailored from several nanometers to hundred nanometers by controlling the etching time and temperature. The dependence of the pore size on the etching time is illustrated in **Figure 3.1d**. Unlike lithography-based nanomaterials, which may suffer from non-uniform UV and temperature exposure, NPG exhibits a uniform nanoporous structure with identical ligaments and nanopores across the entire substrate (**Figure 3.1c**). Such high uniformity and broad coverage of nanopores contribute to robust fluorescence detection for ensemble measurements as well as single QD imaging.

We demonstrated the fluorescence enhancing properties of NPG on two QD species, QD605 (peak emission at 605 nm) and QD525 (peak emission at 525 nm). Both QD species are conjugated with streptavidin and have been widely used in biomolecular sensing and diagnostics. Here, the protein coating additionally serves as a spacer to prevent metal induced quenching by NPG. (66) The distance between QDs and NPG, given by streptavidin, is ~5 nm which falls into the optimal spacer distance (~5-10 nm) for best fluorescence enhancements without obvious quench effect. We used TEM to examine the deposition of QDs on the porous surface of NPG after incubation with a 2 nM QD solution (**Figure 3.2)**. Individual QDs were homogeneously

deposited over the porous structure without detectable aggregation. The semiconductor core of QD605 exhibits an oval shape with a size of 5 nm × 12 nm while QD525 exhibits a round shape with a diameter of 5 nm.

Fluorescence enhancement was first evaluated at the single QD level. **Figure 3.3a** and **d** show fluorescence images of QDs on NPG taken by a basic far-field fluorescence microscope equipped with a CCD camera. Real-time imaging of the QDs reveals the blinking characteristics of single QD fluorescence (**Figure 3.3b** and **e**). The fluorescence intensity of individual QDs was determined by the difference in signal level between the "on" and "off" states. It is worth noting that all the single QD signals observed in this study have a constant fluorescence intensity at the "on" and "off" states. This simply rules out the possibility that the individual bright fluorescence spots come from two or three different QDs. **Figure 3.3c** and **f** are box plots of fluorescence intensity of individual QDs on NPG substrates with different pore sizes. The result suggests that there is an optimum NPG pore size for each of the QDs. The best enhancement for QD605 was seen from the NPG substrate with a 51 nm pore size while QD525 was best enhanced by a pore size of 38 nm. The phenomenon was also confirmed by the fluorescence intensity histogram of QD605 and QD525 on NPG (see **Figure S3.1** in Supporting Information). Similar to previous studies on single molecule fluorescence enhancements,(77, 82) a large variation in fluorescence intensity

of individual QDs was observed with respect to a specific pore size (**Figure 3.3c** and **f**), which was attributed to the random circumstance of individual QDs on the NPG substrate with irregular shapes of nanopores. Because polarized excitation light is used in the fluorescence imaging in this study, only QDs with dipoles well aligned with the polarization of excitation light are best excited while QDs with dipoles vertical to the excitation light polarization cannot be excited regardless of fluorescence enhancement effects. Additionally, the random nanoscopic structure of the NPG film can also induce intensity variation of the LSPR and lead to diverse orientation and distance of QDs to the NPG surface, further contributing to the variation in fluorescence intensity of individual QDs.

We then conducted QD imaging on flat gold substrate without nanopores as a control. A much weaker fluorescent signal was detected because of a very weak fluorescence enhancement for QD525 and an obvious quench effect for QD605 by the non-porous substrate. This weak signal necessitated an increased exposure time of beyond 10's milliseconds to image QDs but the resulting poor temporal resolution confounded observation of the blinking phenomenon for verifying the single dot events. This result demonstrates that the nanoporous structure of NPG plays a vital role for MEF that facilitates robust and quality imaging of single QDs.

In order to more quantitatively evaluate QD fluorescence enhancement by NPG, we deposited a relatively high concentration of QDs (2 nM) on NPG substrates and measured the ensemble fluorescence spectra. We verified through SEM energy disperse spectroscopy (EDS) that a similar number of QDs were deposited on the NPG substrates regardless of pore sizes (see **Figure S3.2** in Supporting Information) and these QDs were uniformly distributed on NPG films (see **Figure S3.3** in Supporting Information). A micro-Raman spectrometer (Renishaw InVia RM 1000) with an excitation laser wavelength of 514.5 nm was used for fluorescence spectra measurements, and the laser power was set at a low value of 0.3 mW with a laser spot of ~5 μm in diameter. As shown in **Figure 3.4a** and **b**, the fluorescence intensities of QDs on NPG were consistently higher than those on the flat gold substrate and glass slide. Ensemble fluorescence intensity is determined by the height of the 605 nm emission peak for QD605 and the 525 nm emission peak for QD525 in individual spectra. The fluorescence enhancement factor $E_f(d)$, as a function of NPG pore size d, was calculated by dividing ensemble QD fluorescence intensity on the NPG with pore size d, $F_{NPG}(d)$, by ensemble QD fluorescence intensity on glass slide, F_{glass}:

$$E_f(d) = F_{NPG}(d) / F_{glass}$$

Once again, we confirmed that the QD fluorescence enhancement factor is critically dependent on the NPG pore size d. The highest fluorescence

enhancements of QD605 and QD525 were achieved from the 51 nm and 38 nm NPG substrates, respectively (**Figure 3.4c** and **d**), which is consistent with single QD imaging results (**Figure 3.3**). The maximum fluorescence enhancement factor of QD525 on the 38 nm NPG substrate is as high as ~100 fold, which, to our knowledge, is the highest reported so far. For QD605, the maximum fluorescence enhancement factor of ~50 fold was observed with the 50 nm NPG substrate.

Both results of single QD fluorescence imaging and ensemble fluorescence spectroscopy suggest that (1) there exists an optimum pore size at which the QD is best enhanced and that (2) the optimum NPG pore size is dependent on the QD species. The mechanism underlying these observations is unclear but is likely attributable to multiple factors. It is known that the spectrum of LSPR induced by NPG strongly depends on its pore size. When water is used as the dielectric medium, the peak excitation of NPG can shift from 500 nm to 650 nm as the pore size increases whereas the LSPR intensity decreases as the nanopore size increases.(83) This red-shifted excitation of NPG may be related to the shift of the optimum pore size from 38 nm for QD525 to 51 nm for QD605. To achieve the best MEF enhancement, the emission wavelength of QD525 and QD605 should be close to the LSPR peaks of NPG in order to enable optimal coupling. Also, compared with isolated gold nanoparticles, bicontinuous NPG, which is an ideally coupled

system of localized and propagating surface plasmon polarizations (SPPs), can induce strongly enhanced local fields from the nearby ligaments and scatter efficiently without any loss of energy. The highly enhanced scattering fields can be absorbed by QDs, leading to highly amplified fluorescence.(69) Besides the double coupling effects, the QD extinction coefficient and hydrodynamic radius can also have a significant impact on overall fluorescence. In particular, since NPG has a 3D porous structure, a high level of enhancement also requires an appropriate geometric match between the QDs and nanopore channels. On the one hand, a stronger electromagnetic field is concentrated in smaller pores of the NPG,(81) hence contributing to a higher extinction coefficient of the QD within the pores. On the other hand, NPG with larger pores has a higher probability of accommodating QDs in the pores. The combination of the two mutually counteracting effects can also lead to an optimum pore size at which the QD fluorescence is best enhanced. Likewise, since the physical size of QD525 (~13 nm) is smaller than that of QD605 (~13 nm × 20 nm), it is reasonable that the optimum pore size for QD525 was smaller than the QD605 counterpart.

In this study we investigated the fluorescence enhancement of QDs by NPG using both single particle imaging and ensemble fluorescence spectroscopic approaches. Compared to non-porous gold substrates, NPG dramatically improves the signal intensity of single QDs. With ensemble

fluorescence measurements, we observed 100-fold and 50-fold enhancements of fluorescence emitted from QD525 and QD605, respectively. Such high levels of enhancement are enabled by both the local surface plasmon resonance and propagating surface plasmon phenomena intrinsic to NPG substrates. In addition, we also observed a strong dependence of fluorescence enhancement on the characteristic pore size of NPG. QD525 and QD605 are best enhanced by the 38 nm and 51 nm NPG substrates, respectively. Furthermore, NPG permits comprehensive coupling between QDs and nanopore channels over a broad substrate area, offering a simple and robust material for both single particle and ensemble fluorescence applications.

3.3. Experimental section

Sample preparation: 700 nm thick NPG films with various pore sizes were prepared by selective dissolution of silver from $Ag_{65}Au_{35}$ (at. %) alloy leaves using 71% nitric acid at room temperature for different time.(81) After carefully washing with distilled water (18.2MΩcm), the as-prepared NPG films with an area of 2 mm × 2 mm were immersed in the QD solution for 1 hour, and then used for measurement after washing with distilled water.

Materials and equipments: CdSe/ZnS core-shell nanocrystals, QD605 (peak emission at 605 nm) and QD525 (peak emission at 525 nm) that were conjugated with streptavidin were purchased from Invitrogen™. Microstructure characterization and QD quantification were performed by

using a scanning electron microscope (SEM, JEOL JIB-4600F) and a transmission electron microscope (TEM, JEOL JEM-2100F). Zeiss LSM 510 Meta confocal microscope was used for single QD imaging, and a micro-Raman spectrometer (Renishaw InVia RM 1000) with incident wavelength of 514.5nm was used for ensemble QD fluorescence measurement.

3.4. Figures

Figure 3.1. Fabrication and microstructure of dealloyed nanoporous gold.

(a) Schematic of the fabrication process of NPG substrates (b) Photographs and microstructure of as-prepared NPG films. (c) 3D electron tomographic image of a representative NPG film. (d) Tunable ligament and pore sizes as a function of etching time. The insets are the top-view SEM micrographs of NPG films with different pore sizes.

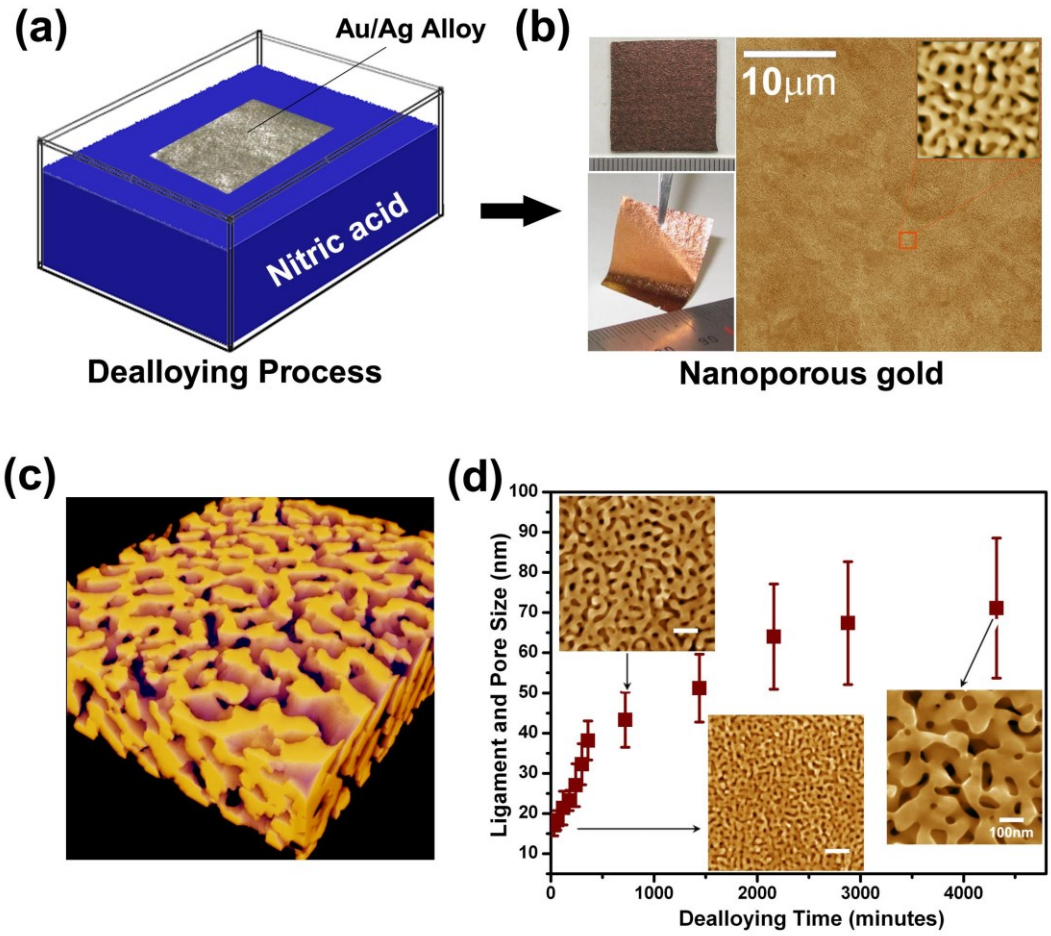

(a)
Au/Ag Alloy
Nitric acid
Dealloying Process
(b)
10μm
Nanoporous gold
(c)
(d)
Ligament and Pore Size (nm)
Dealloying Time (minutes)
100nm

Figure 3.2. TEM images of quantum dots assembled in the nanopore channels of NPG.

(a) QD605; and (b) QD525. 2nM QD aqueous solution is used in the experiment.

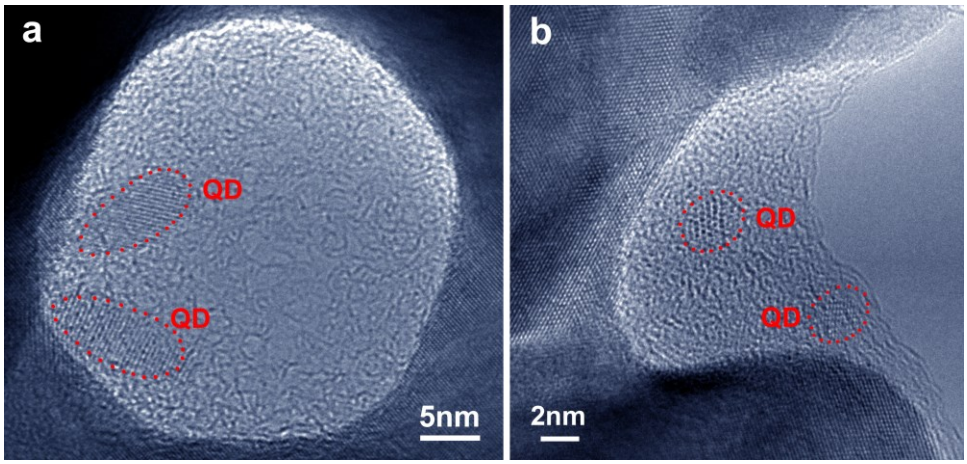

a
QD
QD
5nm
b
QD
QD
2nm

Figure 3.3. Measurements of fluorescence enhancement of single QDs as a function of pore sizes of NPG.

(a) Fluorescence image of QD605 on NPG. (b) A 10ms trace of single QD605 fluorescence. The sudden intensity change (on-off) as a result of fluorescent blinking is characteristic of single QD fluorescence. (c) Box plot of single QD605 fluorescence intensity on NPG of various pore sizes. The fluorescence intensity of each QD is determined as the difference between "on" and "off" states. (d) Fluorescence image of QD525 on NPG. (e) A 10ms trace of single QD525 fluorescence. (f) Box plot of single QD525 fluorescence intensity on NPG of various pore sizes. 20pM QD aqueous solution is used in the experiment.

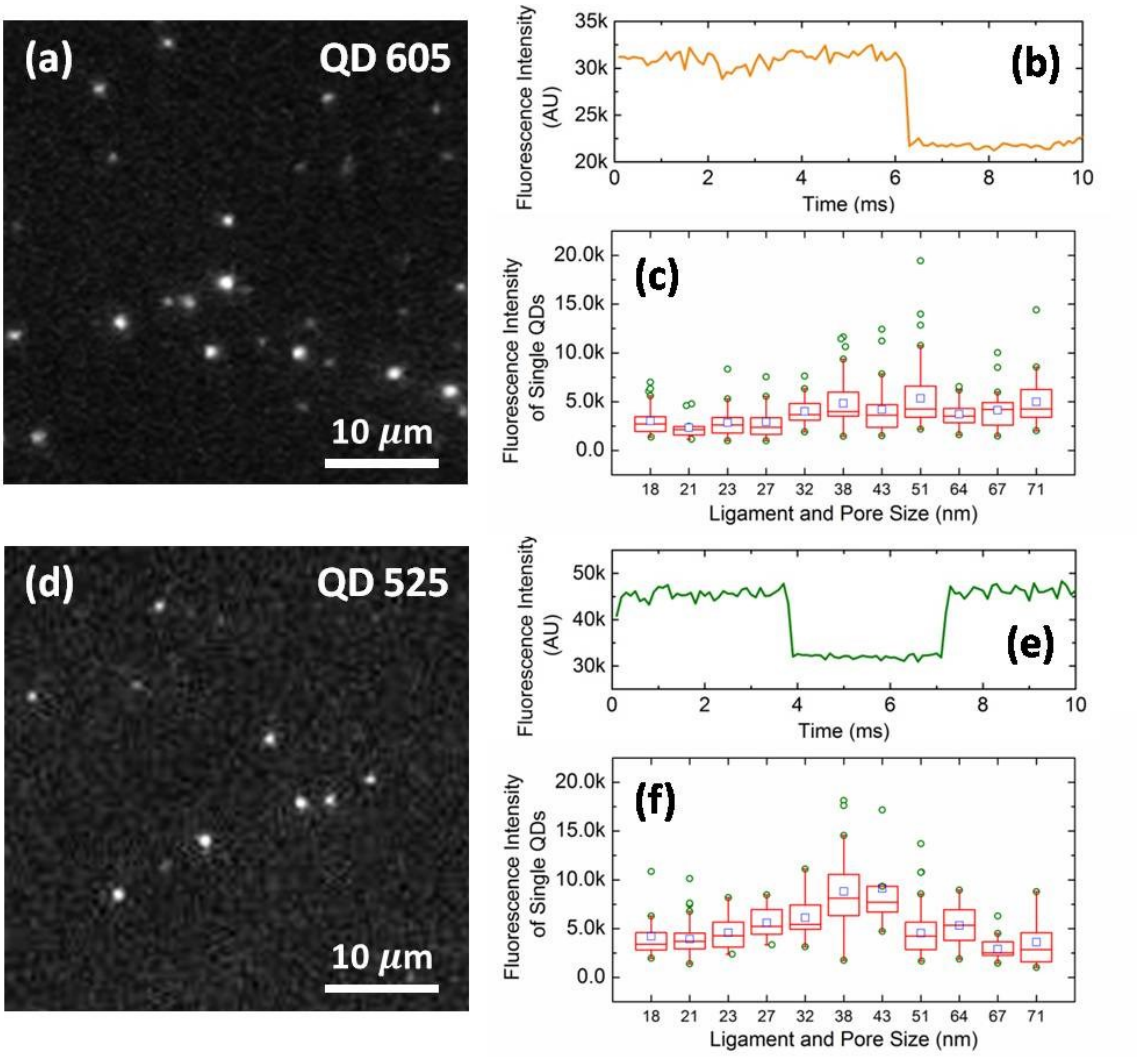
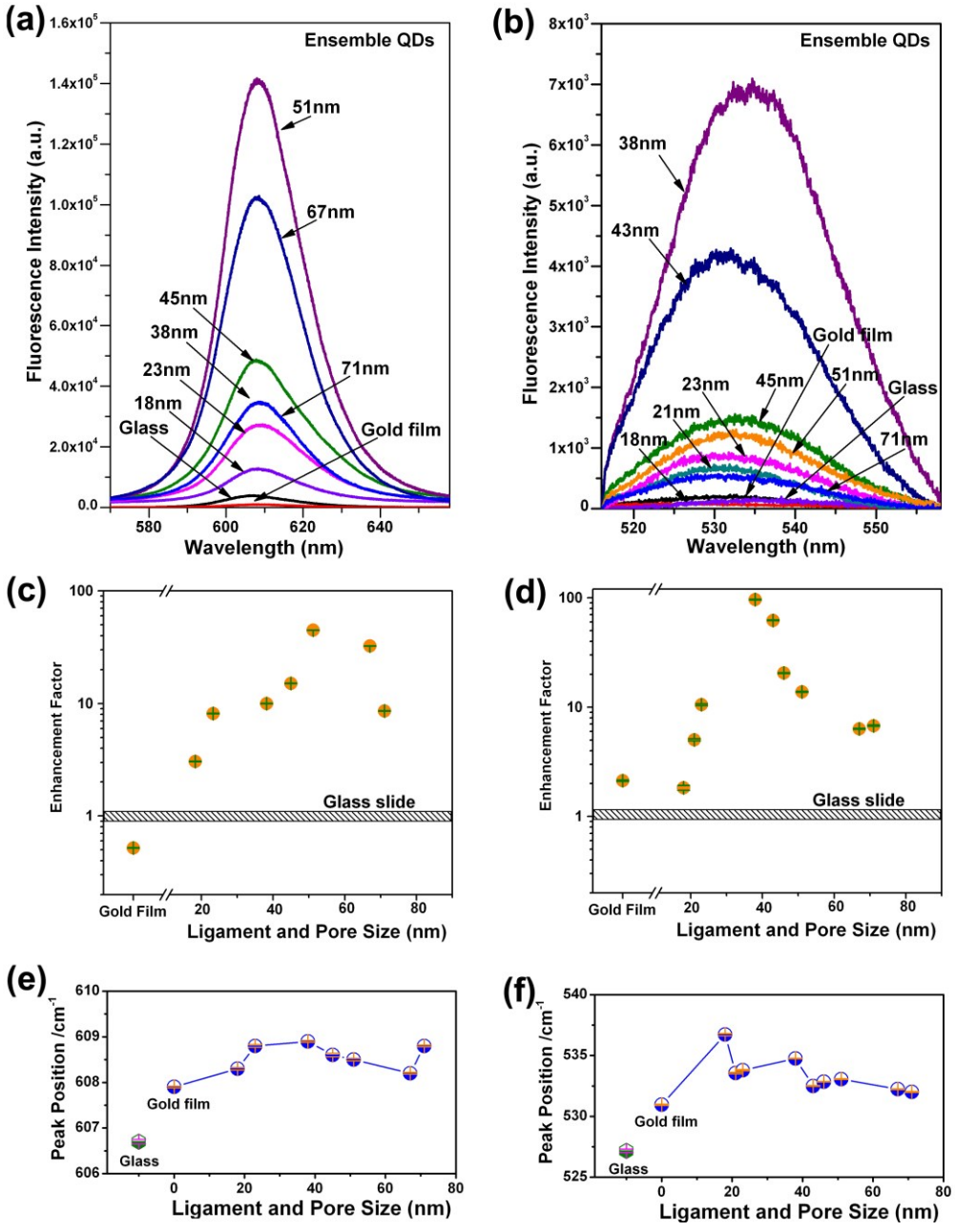

(a)
QD 605
10 μm
(b)
Fluorescence Intensity (AU)
35k
30k
25k
20k
0
2
4
6
8
10
Time (ms)
(c)
Fluorescence Intensity of Single QDs
20.0k
15.0k
10.0k
5.0k
0.0
18
21
23
27
32
38
43
51
64
67
71
Ligament and Pore Size (nm)
(d)
QD 525
10 μm
(e)
Fluorescence Intensity (AU)
50k
40k
30k
0
2
4
6
8
10
Time (ms)
(f)
Fluorescence Intensity of Single QDs
20.0k
15.0k
10.0k
5.0k
0.0
18
21
23
27
32
38
43
51
64
67
71
Ligament and Pore Size (nm)

Figure 3.4. Ensemble fluorescence enhancement.

Fluorescence spectra of QD605 (a) and QD525 (b) on glass slide, flat gold and the NPG films. (c) and (d) show the QD605 and QD525 fluorescence enhancements by the NPG substrates, respectively, as a function of the pore sizes. The enhancement factors were determined as the ratio of the peak intensity of QDs on the NPG and the glass slide. (e) and (f) show the emission peak position of QD605 and QD525, respectively, on the NPG, non-porous gold substrate and glass slide. 2nM QD aqueous solution is used in the experiment.


(a)
Ensemble QDs
51nm
67nm
45nm
38nm
23nm
18nm
Glass
71nm
Gold film
Fluorescence Intensity (a.u.)
Wavelength (nm)
(b)
Ensemble QDs
38nm
43nm
Gold film
23nm
45nm
51nm
Glass
21nm
18nm
71nm
Fluorescence Intensity (a.u.)
Wavelength (nm)
(c)
Enhancement Factor
Glass slide
Gold Film
Ligament and Pore Size (nm)
(d)
Enhancement Factor
Glass slide
Gold Film
Ligament and Pore Size (nm)
(e)
Peak Position /cm-1
Gold film
Glass
Ligament and Pore Size (nm)
(f)
Peak Position /cm-1
Gold film
Glass
Ligament and Pore Size (nm)

3.5. Supplementary Information

Figure S3.1. Histogram of single (a) QD525 and (b) QD605 fluorescence on gold film and NPG films with different ligament sizes. (The back fluorescence was subtracted from the entire fluorescence intensity. 20pM QD aqueous solution is used in the experiment.

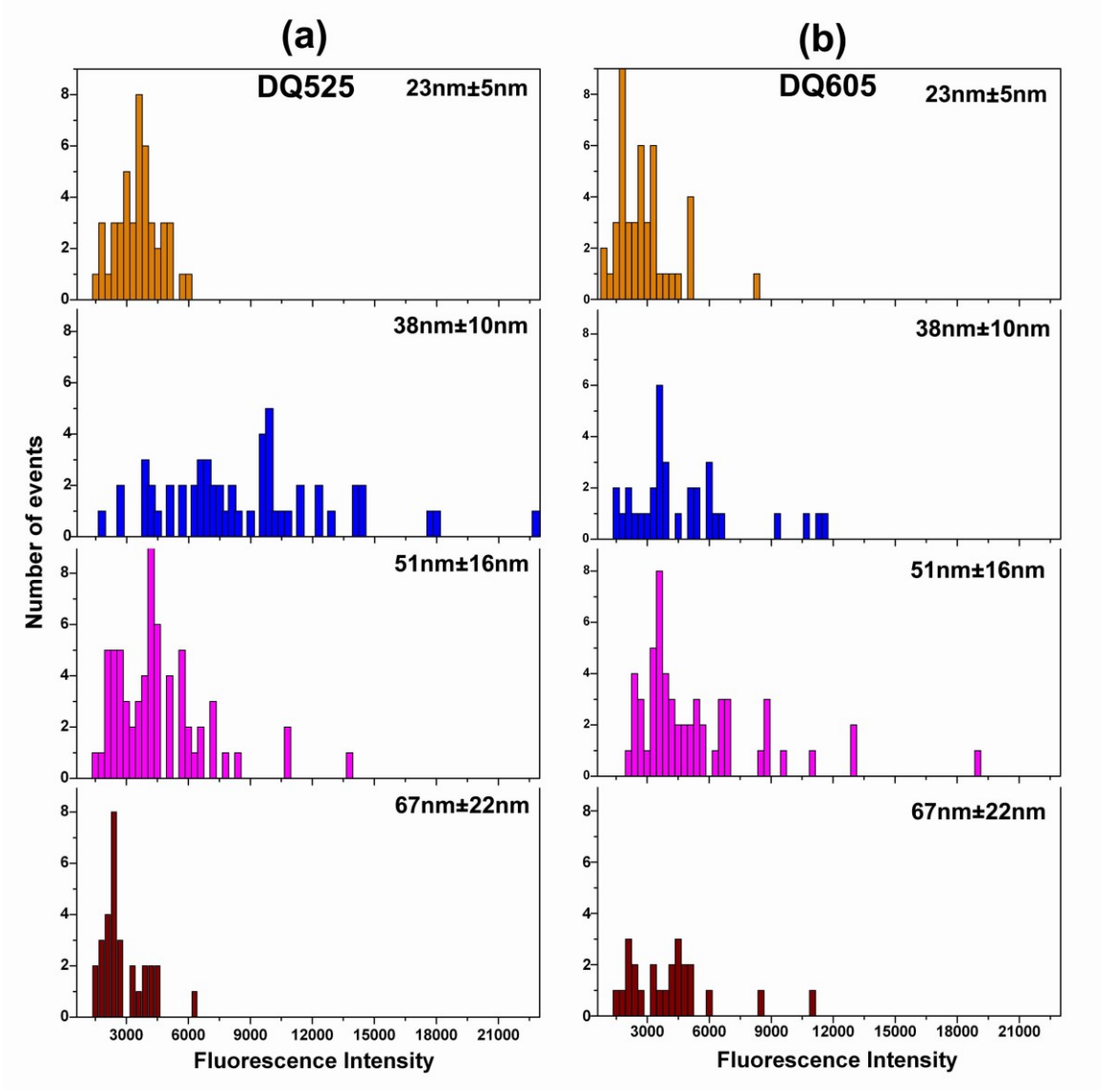

(a)
DQ525
23nm±5nm
38nm±10nm
51nm±16nm
67nm±22nm
Number of events
Fluorescence Intensity
(b)
DQ605
23nm±5nm
38nm±10nm
51nm±16nm
67nm±22nm
Fluorescence Intensity

Figure S3.2. Energy Disperse Spectroscopy (EDS) analysis of QD605 decorated NPG films. The ratio between Zn and Au&Ag is shown in green dots, and the red line is drawn for reference. The blue dot shows the relative QD amount on glass slide which is evaluated by compare the QD on gold thin film coated glass and gold thin film covered QD on glass slide. 2nM QD aqueous solution is used in the experiment.

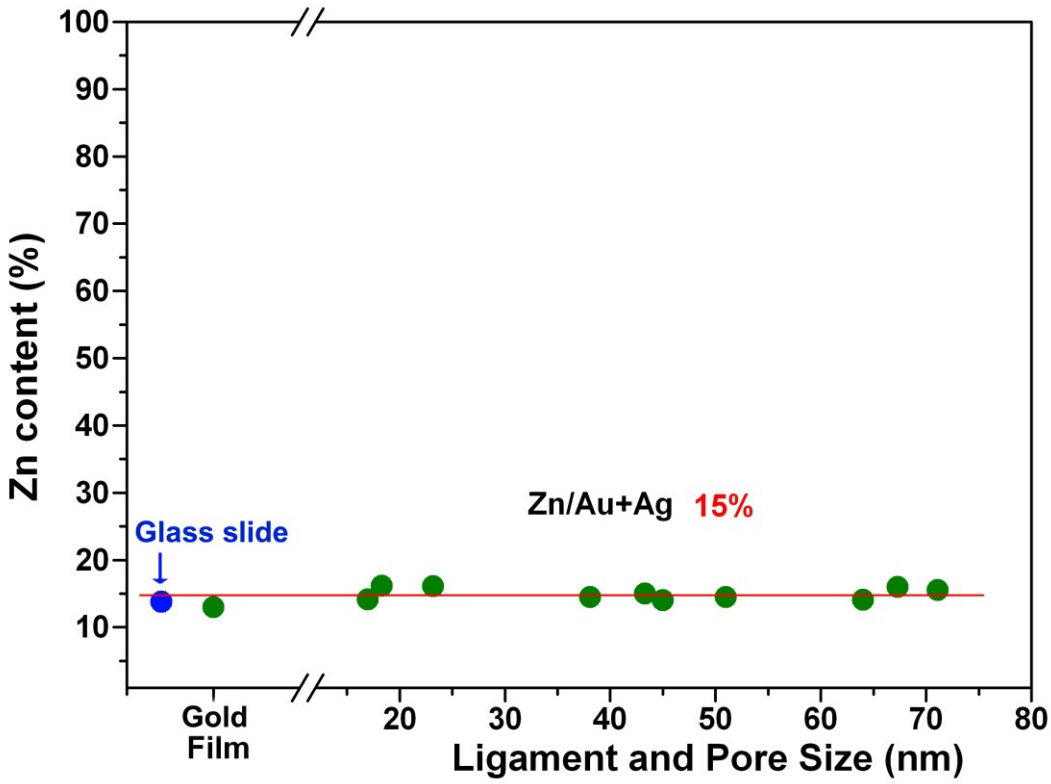
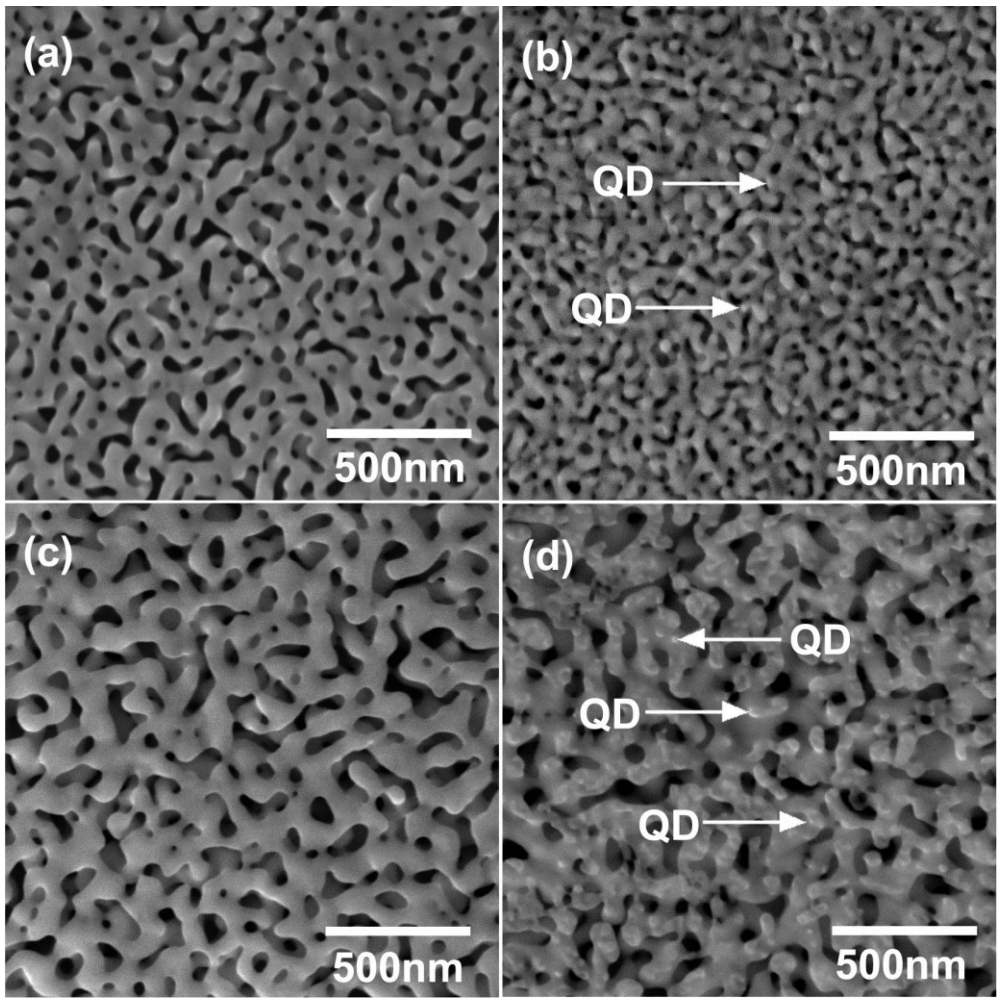

Zn content (%)
100
90
80
70
60
50
40
30
20
10
Zn/Au+Ag 15%
Glass slide
Gold
Film
20
30
40
50
60
70
80
Ligament and Pore Size (nm)

Figure S3.3. Micrographs showing representative morphological of NPG flims without and with QDs coating. SEM images of 38nm-NPG films (a) without QDs and (b) with QD525. SEM images of 51nm-NPG films (c) without QDs and (d) with QD605. 2nM QD aqueous solution is used in the experiment.


(a)
500nm
(b)
QD
QD
500nm
(c)
500nm
(d)
QD
QD
QD
500nm

Chapter 4: Preparing Pre-adenylated Adapter

Song,Y., Liu,K.J. and Wang,T.H. Efficient synthesis of stable adenylated DNA and RNA adapters for microRNA capture using T4 RNA Ligase 1. In Preparation.

4.1. Introduction

Since the discovery of small RNAs and their significance in various biological activities, a number of assays have been invented to specifically quantify the microRNA expression of various biological samples. MicroRNA, an important class of the small RNAs, are short RNAs that regulates post-transcriptional gene regulation. In general, widely adopted assays such as Reverse Transcription-quantitative Polymerase Chain Reaction (RT-qPCR) and microarray require microRNA modification methods. Since microRNAs are too short to manipulate as they are, most modification methods elongate microRNAs by either addition of poly(A) tail or ligation of an oligonucleotide strand called "adapter". Among these methods, adapter ligation method is most versatile and has been integrated into many microRNA expression profiling assays.

Typical mature microRNAs are as short as 17 to 24 nt and have 5'-phosphate and 3'-hydroxyl termini. Their short length and 5' and 3' functionalities render it susceptible to intra- and intermolecular ligation during the microRNA – adapter ligation step. For example, 5' and 3' termini

of the microRNA can be ligated together, forming a circularized product which cannot be further ligated to an adapter. Alternatively, microRNAs can be ligated to other DNA or RNA molecules within the sample instead of adapters. In both cases, the side reactions make it elusive to acquire the most accurate picture of the microRNA expression profile. A predominant method to circumvent the problem has been to ligate a pre-adenylated adapter to microRNA molecules using a mutant of T4 RNA Ligase 2 which lacks the adenylation capability. Under this method, ligation reaction only happens between RNA and "already activated" pre-adenylated adapter molecules, but not between target molecules themselves. There has been multiple studies devoted for achieving efficient and unbiased adapter ligation to the microRNA using various RNA and DNA ligases.(84, 85, 86, 87)

Highly efficient microRNA capture is only feasible if pre-adenylated adapters could be synthesized in a large quantity and with little impurities. The adapter pre-adenylation was first achieved by chemical synthesis, in which AMP was activated and adenosine 5'-phosphorimidazolidate was coupled to the 5'-phosphorylated oligonucleotide, followed by purification of the pre-adenylated adapter from unreacted substrates.(88) Then an enzymatic method was invented so as to achieve a higher yield and to reduce reaction complexity. In this method, the adapter was first annealed to the template, adenylated by template-dependent T4 DNA Ligase, and then

purified from the template.(89) T4 DNA Ligase 1 catalyzes phosphodiester bond formation between the phosphorylated 5' terminal and the 3' hydroxyl terminal through three steps.(90, 91) The ligase first forms a covalent bond with ATP, generating a ligase-(lysyl-N)–AMP intermediate and releasing a pyrophosphate. Next, adenylylated RNA (AppRNA) is formed as AMP is transferred from the ligase to the 5'-phosphate of RNA. The enzyme lacks the specificity to distinguish RNA from DNA; hence this second step reaction occurs for both RNA and DNA oligonucleotides. Finally, the ligase catalyzes the 3'-5' phosphodiester bond formation via an attack by an RNA 3'-OH on the AppRNA or AppDNA, releasing AMP. If there is no free 3'-OH terminal available in the solution, the enzyme eventually leaves the adenylated RNA or DNA, which can be used as pre-adenylated adapter for microRNA capture. In all of chemical and enzymatic methods above, the laborious purification step is a practical obstacle from user's perspective. A template-independent enzymatic method has also been proposed using thermostable Archaeal RNA Ligase, which was originally discovered as a ligase that catalyzes intramolecular ligation of single-stranded RNA and DNA.(92) This again raises fears that Archaeal RNA Ligase carried over from adenylation step could cause microRNA circularization in the later microRNA capture step, hindering accurate microRNA quantification as a result. Furthermore, while the enzyme adenylates the adapter at optimum temperature of 65 ºC, the reversal reaction prevails at lower temperature of 25 ºC, which is the

optimum temperature for T4 RNA Ligase 2 truncated used in microRNA – adapter ligation reaction, hence reducing the amount of usable adapter in the solution.(93)

Herein we report a simple single step, purification-free method which allows efficient and irreversible adenylation of DNA and RNA adapters using T4 RNA Ligase 1. The enzyme can be easily heat-inactivated, posing no complication of microRNA circularization or adapter de-adenylation in the subsequent microRNA capture step using T4 RNA Ligase 2 truncated. The pre-adenylated adapter has been verified to capture microRNAs within total RNA samples in a highly efficient and unbiased manner using the protocol previously reported in our group.(87)

4.2. Results

4.2.1. Adenylation bias due to adapter

In order to establish an efficient method to adenylate the adapter without a need for purification, we chose T4 RNA Ligase 1, which catalyses the ligation between DNA or RNA single strands. We hypothesized that the RNA ligase conjugates ATP to RNA more efficiently than DNA. In order to test the hypothesis, we compared the adenylation efficiency of the enzyme using 5'-phosphorylated DNA (DNA-A) and a RNA/DNA hybrid (RNA-A) molecules (**Table S1**). While the enzyme did adenylate the RNA/DNA hybrid more efficiently at low PEG levels, the difference was unnoticeable at high PEG

levels, in which case both adapters were adenylated efficiently (**Figure 4.1**). We further questioned whether the first nucleotide at the 5' terminal of the adapter makes substantial difference in adenylation efficiency. To emphasize the difference between adapters, we added large amount of adapter into the reaction (10 μM). Under our initial conditions of 10 μM adapter, 1 mM ATP, and 100 unit ligase per 1 nanomole adapter, and 1 hour reaction time, the adapters that had cytosine and guanine nucleotide at the 5' terminal were the most and least efficiently adenylated, respectively (**Figure S4.1**), with other adapters in-between (data not shown). The adenylation efficiency increased as more PEG was added for all adapters tested; however, the critical PEG level, at which the adenylation efficiency drastically increases, was different depending on the nucleotide at the 5' terminal. While RNA-C adapter was well adenylated at PEG level as low as 5%, DNA-G was much harder to adenylate even at 30% PEG level. While an increase of PEG level from 20% to 30% enhanced adenylation efficiency significantly for RNA-A adapter, the PEG increase hardly meant anything for RNA-C adapter, which is already completely adenylated at such PEG levels. Whereas the efficiency of the microRNA – adapter ligation reaction using T4 RNA Ligase 2 truncated decreases at excessively PEG levels(87), a decrease of adenylation efficiency using T4 RNA Ligase 1 was not observed as PEG level increased up to 35%. Since no drawback was observed at high PEG levels, we conclude

that 35% could be used as a versatile PEG level for all DNA and RNA adapters.

4.2.2. Optimization of enzyme-to-adapter ratio, ATP concentration, adapter concentration and time

Next we examined the effects of enzyme/substrate ratio as well as ATP concentration on adenylation reaction. Enzyme/substrate ratio should be optimized to keep cost low but still yield sufficiently large amount of adenylated adapter. ATP should be close to stoichiometric so that carry-through is minimized while not reducing adenylation efficiency. It was found that 300 unit ligase per 1 nanomole adapter, combined with 1 mM ATP, generated highly efficient adenylated adapters within 1 hour (**Figure 4.2**).

In order to avoid unnecessary dilution at the stage of microRNA capture, adapter solution should be prepared as a highly concentrated solution. To examine how much adapter can be adenylated within 1 hour, we changed the adapter concentration in the solution while keeping the enzyme amount the same. Up to 2 uM of adapter was fully adenylated within 1 hour, again corresponding to 300 unit ligase per 1 nanomole adapter (**Figure 4.3**). With longer incubation time (e.g., over-night), DNA-A adapter of 10 uM could be fully adenylated (data now shown).

For previous reactions, the solution was incubated at 37 °C for 1 hour. We questioned whether this incubation time was enough. As shown in

Figure 4.4, the adapter was found to be efficiently adenylated by incubating for as short as 30 minute under optimized 300 unit ligase to 1 nanomole adapter and 1 mM ATP condition.

Under the above-described optimized conditions (35% PEG; 1 mM ATP; 2 uM adapter; 300 unit ligase per 1 nanomole adapter; 1 hour reaction time), all of RNA-A, RNA-T, RNA-C, DNA-A, DNA-T, and DNA-C adapters generated perfectly adenylated adapters within 1 hour (**Figure 4.5**), whereas DNA-G and RNA-G adapters required longer incubation time of 120 minutes (**Figure S4.2**).

4.2.3. Verification of the adenylated adapter

The validity of the adenylated adapter was verified in microRNA – adapter ligation reaction (**Figure 4.6**). The DNA-A adapter was first adenylated using T4 RNA Ligase 1 followed by heat-inactivation of 65°C for 15 min. The adapter was first used to capture synthetic let-7a microRNA that was labeled with Cy3 at its 5' terminal. The capture efficiency, defined as (Cy3 signal of captured let-7a band) / (the Cy3 signal of captured let-7a band + the Cy3 signal of uncaptured let-7a band), was 91%. In addition to synthetic let-7a microRNA, the adapter was used to capture RNA molecules within 100 ng of the total RNA extracted from human pancreatic tissues. The RNA molecules within the extracted total RNA sample was successfully captured by pre-adenylated, Cy5-labeled adapter as evidenced by the bands in high molecular

weight region of the gel. These are likely the ligation products of rRNA and the adapter. As a further verification, let-7a microRNA was spiked into the total RNA sample and captured by the pre-adenylated adapter using our previously described method. In this case, the capture efficiency of the let-7a microRNA was 70%. At the same time, large RNA molecules within the total RNA were also captured by the adapter as evidenced by the ligation product bands in high molecular region in Cy5 channel. This indicates that the pre-adenylated adapter prepared with the proposed method is fully capable of capturing microRNA molecules within the total RNA sample in an efficient way.

As another verification of the adapter functionality, total RNA was captured while being titrated from 1,000 ng to 10 ng (**Figure S4.3**). In addition to different amount of total RNA, each reaction also included constant 10 nM of six microRNA (miR-let-7a, 16, 21, 26a-1, 29b and 34a) spiked-in. The ligation products in high molecular region in the top, again representing total RNA-derived large RNA molecules captured by the adapter, were linearly proportional to the amount of total RNA input. No complication from synthetic RNA was observed. This demonstrates that the adenylated adapter is able to capture each RNA species proportionally to its presence within the total RNA, enabling accurate downstream quantification.

4.2.4. Side Product Ligation

We compared both the functionality and the side effects of pre-adenylated adapters synthesized by T4 RNA Ligase 1 and Archaeal RNA Ligase. DNA-A adapter was adenylated using either T4 RNA Ligase 1, using our proposed method, or Archaeal RNA Ligase, following the recommended protocol (https://www.neb.com). The adapter was adenylated at high efficiency under both protocols and RNAs were well captured using our previously described methods (Figure 4.6 and **Figure S4.4**).(87) We conclude that there was no substantial difference between the two methods in term of adenylation efficiency and utility in the microRNA – adapter ligation reaction.

Next, we examined the side effects of both methods. As previously described by multiple groups(92, 93), thermostable Archaeal RNA Ligase adenylates adapter at 65 ºC but the reverse (de-adenylation) reaction prevails at non-optimum temperature such as 37 ºC. Indeed, the adapter adenylation was observed during microRNA – adapter ligation reaction by Zhelkovsky, et. al. Under our conditions, however, adapter de-adenylation was observed neither for T4 RNA Ligase 1-prepared adapter nor for Archaeal RNA Ligase-prepared adapter (Figure 4.6 and S4). Nevertheless, the other side reaction – RNA self-circularization – was observed only when Archaeal RNA Ligase-prepared adapters were used for microRNA capture (**Figure S4.5**). In this experiment, a mixture of (1) a single stranded DNA mimicing 80 nt-long let-

7a precursor sequence that was labelled with 5'-phosphate and 3'-OH and (2) the Cy3-labeled let-7a microRNA were used as target. The mixture was then incubated with pre-adenylated adapter synthesized by either method. The post-microRNA capture sample was analysed using denaturing PAGE gel stained with SYBR® Safe DNA Gel Stain (Life Technologies, Carlsbad, CA). While the 80 nt-long DNA was long enough to be properly stained by the intercalating fluorescence dye, the short let-7a and the adapter were not efficiently visualized on the gel. Only when pre-adenylated adapter synthesized by Archaeal RNA Ligase was used, circularized DNA band was observed on the gel. The amount of the circularized DNA decreased as higher PEG levels were introduced to the reaction; however, the circularization reaction was not entirely quenched even at 25% PEG, the optimum PEG level for microRNA – adapter ligation.

4.3. Discussion

In this report we report a quick, simple and robust way to synthesize the pre-adenylated adapter using T4 RNA Ligase 1 without a need for template or subsequent purification steps. The adenylation is irreversible and the enzyme can be effectively heat-inactivated prior to the microRNA capture step, eliminating the possibility of microRNA circularization.

Polyethylene glycol (PEG) is known to enhance the activity of T4 RNA ligases by molecular crowding.(94) We recently found that optimizing PEG

level allows for efficient and unbiased ligation between microRNA and adapter using T4 RNA Ligase 2 truncated K227Q.(87) In case of microRNA – adapter ligation without total RNA, no more increase in the ligation efficiency was observed after PEG level reached 25%. In presence of total RNA, higher PEG levels could even be detrimental to the reaction efficiency. On contrary, in case of adapter adenylation reaction by T4 RNA Ligase 1, no reduction in reaction efficiency was observed as higher PEG level was introduced to the reaction up to 35%. At 35% PEG level, eventually all adapters tested in our report were adenylated for 100% efficiency.

Many microRNA detection assays are designed to detect microRNA targets from total RNA samples. Since adapter ligation step is non-specific to microRNA, other RNA molecules within the sample such as rRNA, tRNA, mRNAs and other small RNAs can also be captured by adapters. In other words, such "background molecules" compete with microRNA in the ligation reaction, suggesting that microRNAs could be "lost" (i.e., not captured by adapter and would not be detected) unless the pre-adenylated adapter is synthesized and provided in the ligation reaction at a high concentration. We previously described that 200 nM of adenylated adapter is enough to capture microRNAs from 500 ng of total RNA. Under the proposed method, adapter can be efficiently adenylated at sufficiently high concentration (2 uM) with 1

hour of incubation. An even larger amount of adapters can be adenylated if longer incubation time is used.

Although synthesis of pre-adenylated adapter is an essential issue for microRNA – adapter ligation step, most methods proposed so far involves purification of the pre-adenylated adapter from rest of the molecules, including non-adenylated adapters and templates. Purification-free, template independent method using thermostable Archaeal RNA Ligase has been proposed; however, the method has been shown to have generate de-adenylated adapter from fully adenylated adapter, suggesting that the Archaeal RNA Ligase could be partially active in the microRNA – adapter ligation step. This implies another potential drawback caused by the partially active Archaeal RNA Ligase during the microRNA – adapter ligation reaction: the microRNA circularization. We found it worthwhile to verify this possibility because Archaeal RNA Ligase was initially discovered as an enzyme that preferentially catalyses RNA circularization.(92)

To our best knowledge, Archaeal RNA Ligase has never been tested with circularization reactions, as a background reaction of microRNA – adapter ligation, using a DNA or RNA target which is labelled with 5'-phosphate and 3'-OH before. Our experiment using 5' phosphorylated DNA target implies that the Archaeal RNA Ligase could potentially introduce circularized RNA product in the second step. To make matter even worse, Archaeal RNA

Ligase prefers RNA targets to DNA targets, suggesting the circularization reaction could be even more pronounced for short microRNA molecules. The amount of circularized microRNA could be different depending on the condition under which microRNA are captured by the adapter. For instance, less circularized microRNA was formed when high PEG levels were used. It is expected that some microRNAs are more likely to be circularized than others due to their sequence or secondary structure, hence introducing additional bias in microRNA expression profiling. In such cases, adapter adenylated by T4 RNA Ligase 1 could be a safer alternative.

4.4. Methods

Oligonucleotides. Both microRNAs and adapters were synthesized by Integrated DNA Technologies (Coralville, IA). MicroRNAs, whose sequences were taken from miRBase (www.mirbase.org), were labeled with Cy3 and –OH at their 5' and 3' termini, respectively. The adapters were based on modified 3' modban adapters and were labeled with phosphate and Cy5 at their 5' and 3' termini, respectively. All sequences were HPLC purified. The sequences are listed in **Table S1**.

Adenylation Reaction. Unless otherwise indicated, the adenylation reaction was performed in 25 μL solution containing 2 μM DNA-A adapter (Table S1); 1X T4 RNA Ligase Buffer (New England Biolabs, Ipswich, MA); 35% PEG; 1 mM ATP; and 600 unit / mL, or 300 unit ligase per 1 nanomole adapter of T4

RNA Ligase 1 (New England Biolabs, Ipswich, MA). The solution was held at 37 ºC for 1 hour and 65 ºC for 15 min for heat-inactivation of the T4 RNA Ligase 1. Alternatively, adapter pre-adenylation was conducted using 5' DNA Adenylation Kit (New England Biolabs, Ipswich, MA), following the instructions from the manufacture.

microRNA – Adapter Ligation. microRNA was ligated to pre-adenylated adapter using T4 RNA Ligase 2 truncated K227Q using the protocol as previously described.

PAGE Analysis. The samples were analyzed on precast 15% TBE-urea polyacrylamide gels (Bio-Rad, Hercules, CA). 5 µL of sample was mixed with 5 µL of loading buffer and heated for 5 minutes at 95 ºC. The sample was then loaded into the gel and run for 20 min at 300V. The separated gels were scanned using a Typhoon 9410 variable mode imager (GE Healthcare, Piscataway, NJ). The gel images were visualized using ImageQuant (GE Healthcare, Piscataway, NJ) to obtain lane profiles.

4.5. Tables and Figures

Table S4.1. Oligonucleotides

Name	Sequence 5' -> 3'
DNA-A	5' - /5Phos/ACTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
DNA-T	5' - /5Phos/TCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
DNA-G	5' - /5Phos/GCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
DNA-C	5' - /5Phos/CCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
RNA-A	5' - /5Phos/rACTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
RNA-U	5' - /5Phos/rUCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
RNA-G	5' - /5Phos/rGCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
RNA-C	5' - /5Phos/rCCTGTAGGCACCATCAATC/3Cy5Sp/ - 3'
let-7a	/5Cy3/rUrGrArGrGrUrArGrUrArGrGrUrUrGrUrArUrArGrUrU
miR-16	/5Cy3/rUrArGrCrArGrCrArCrGrUrArArArUrArUrUrGrGrCrG
miR-21	/5Cy3/rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA
miR-26a	/5Cy3/rUrUrCrArArGrUrArArUrCrCrArGrGrArUrArGrGrCrU
miR-29b	/5Cy3/rUrArGrCrArCrCrArUrUrUrGrArArArUrCrArGrUrGrUrU
miR-34a	/5Cy3/rUrGrGrCrArGrUrGrUrCrUrUrArGrCrUrGrGrUrUrGrU
let-7a precursor DNA	/5Phos/TGGGATGAGGTAGTAGGTTGTATAGTTTTAGGGTCACA CCCACCACTGGGAGATAACTATACAATCTACTGTCTTTCCTA

Figure 4.1. The comparison of DNA and RNA/DNA hybrid adapter in adenylation reaction under various PEG levels.

Both DNA-A and RNA-A adapters are efficiently adenylated under 35% PEG.

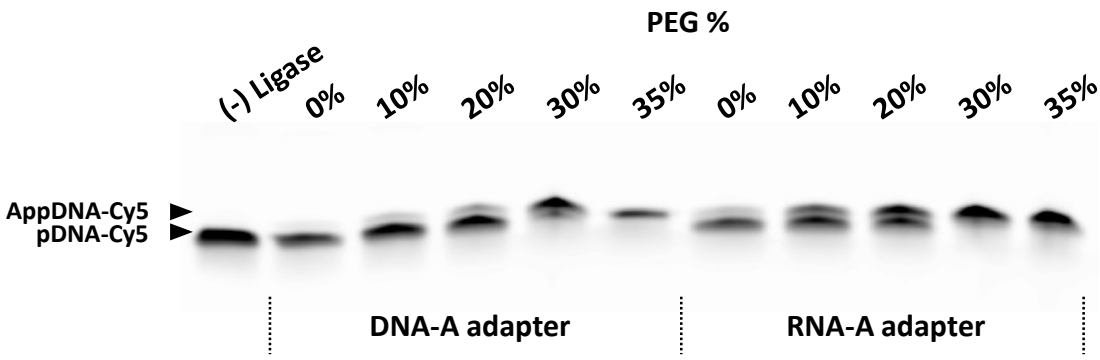

PEG %
(-) Ligase
0%
10%
20%
30%
35%
0%
10%
20%
30%
35%
AppDNA-Cy5
pDNA-Cy5
DNA-A adapter
RNA-A adapter

Figure 4.2. ATP and Enzyme optimization.

A combination of (i) 300 unit ligase per 1 uM adapter and (ii) 250 uM or more of ATP yields high adenylation efficiency.

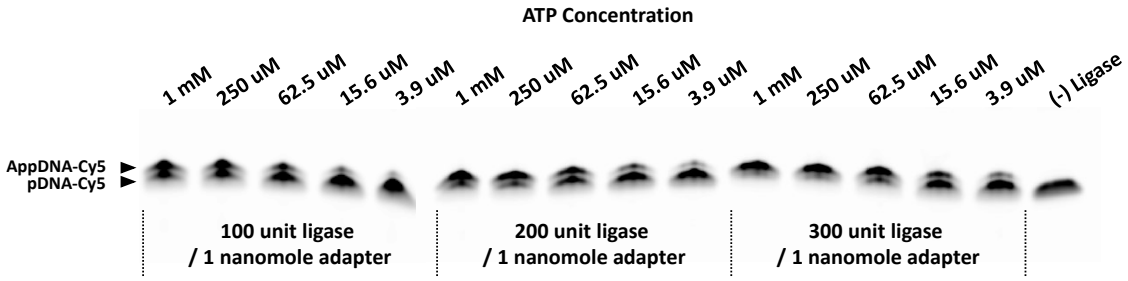

ATP Concentration
1 mM
250 uM
62.5 uM
15.6 uM
3.9 uM
1 mM
250 uM
62.5 uM
15.6 uM
3.9 uM
1 mM
250 uM
62.5 uM
15.6 uM
3.9 uM
(-) Ligase
AppDNA-Cy5
pDNA-Cy5
100 unit ligase
/ 1 nanomole adapter
200 unit ligase
/ 1 nanomole adapter
300 unit ligase
/ 1 nanomole adapter

Figure 4.3. Adenylation of various amount of adapter.

The number in the bracket indicates the ratio of unit ligase per nanomole adapter. As much as 2 uM of adapter can be adenylated in 1 hour.

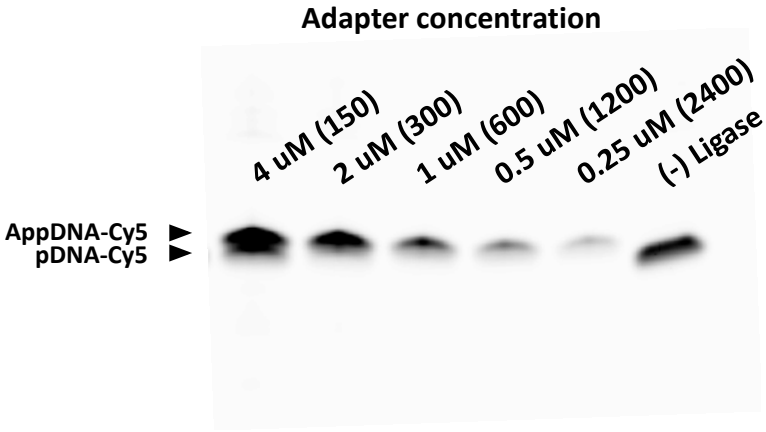

Adapter concentration
4 uM (150)
2 uM (300)
1 uM (600)
0.5 uM (1200)
0.25 uM (2400)
(-) Ligase
AppDNA-Cy5
pDNA-Cy5

Figure 4.4. The yield of adenylation reaction with various incubation time.

Adapter adenylation is complete in less than 30 min for DNA-A adapter.

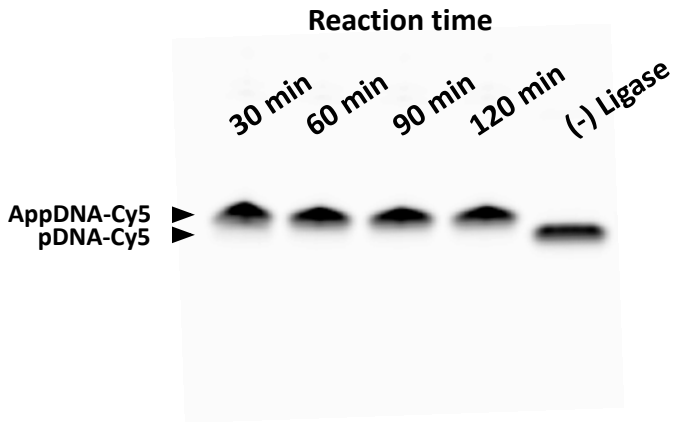

Reaction time
30 min
60 min
90 min
120 min
(-) Ligase
AppDNA-Cy5
pDNA-Cy5

Figure 4.5. Adenylation of DNA and RNA/DNA hybrid adapters.

DNA-G and RNA-G adapters are adenylated less efficiently compared to other adapters within 1 hour.

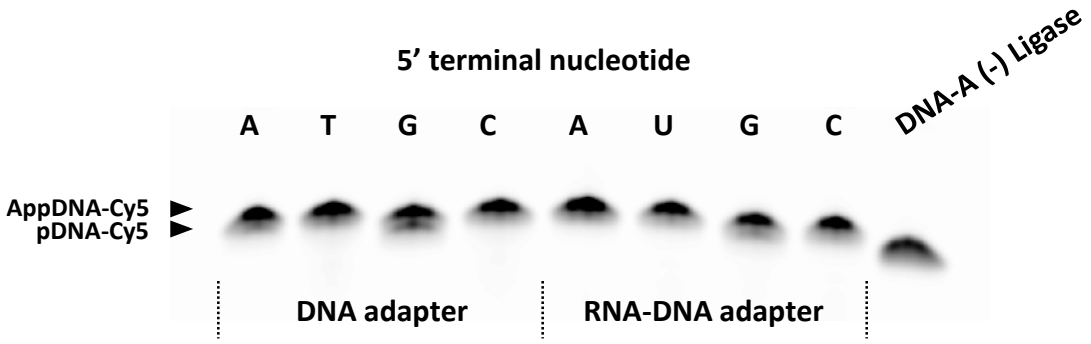
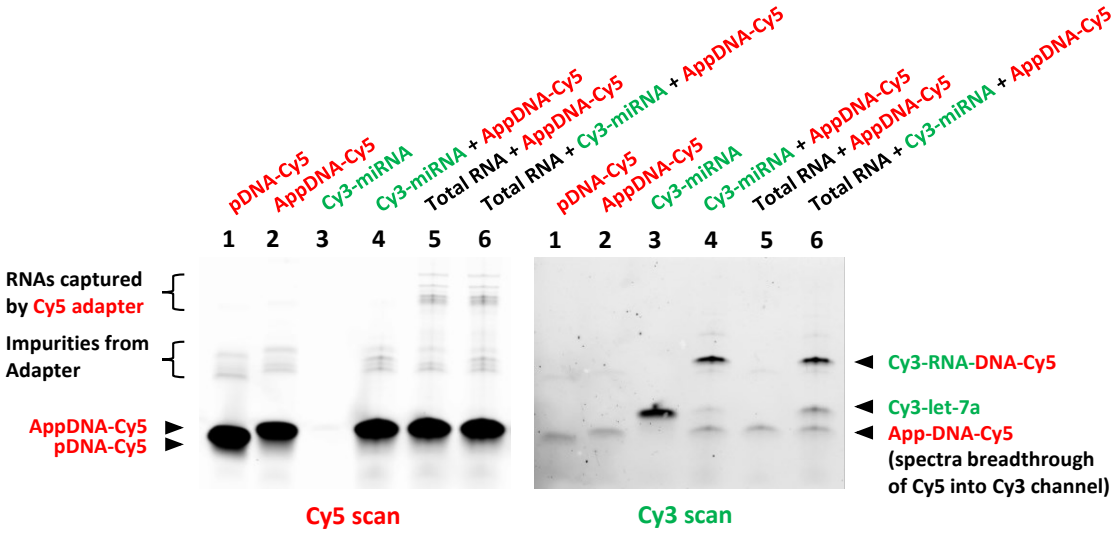


Figure 4.6. Demonstration of microRNA – adapter ligation using pre-adenylated adapter synthesized with T4 RNA Ligase 1.

Both Cy5-spectra scan (excitation at 633 nm and emission at 670 nm) and Cy3-spectra scan (excitation at 532 nm and emission at 580 nm) are displayed. DNA-A adapter is effectively adenylated (lane 1: unadenylated adapter; lane 2: adenylated adapter) using the proposed method. Let-7a microRNA is effectively ligated to the adapter with and without addition of 100 ng of total RNA extracted from human pancreatic tissue (lane 4 and 6, respectively). Large RNA molecules within the total RNA are also captured by the adapter (lane 5). No de-adenylation of the pre-adenylated adapter is observed.


pDNA-Cy5
AppDNA-Cy5
Cy3-miRNA
Cy3-miRNA + AppDNA-Cy5
Total RNA + AppDNA-Cy5
Total RNA + Cy3-miRNA + AppDNA-Cy5
1 2 3 4 5 6
RNAs captured by Cy5 adapter
Impurities from Adapter
AppDNA-Cy5
pDNA-Cy5
Cy5 scan
Cy3 scan
Cy3-RNA-DNA-Cy5
Cy3-let-7a
App-DNA-Cy5
(spectra breadthrough of Cy5 into Cy3 channel)

4.6. Supplementary Data

Supplementary figures follow below.

Figure S4.1. Adenylation of RNA-A, DNA-A, RNA-C and DNA-G adapters with various PEG levels.

The lower ligase input of 100 unit/μM was used. The adapter is adenylated at dramatically different efficiency depending on the first nucleotide of the adapter as well as PEG levels. DNA-G adapter and RNA-C adapters are least and most efficiently adenylated.

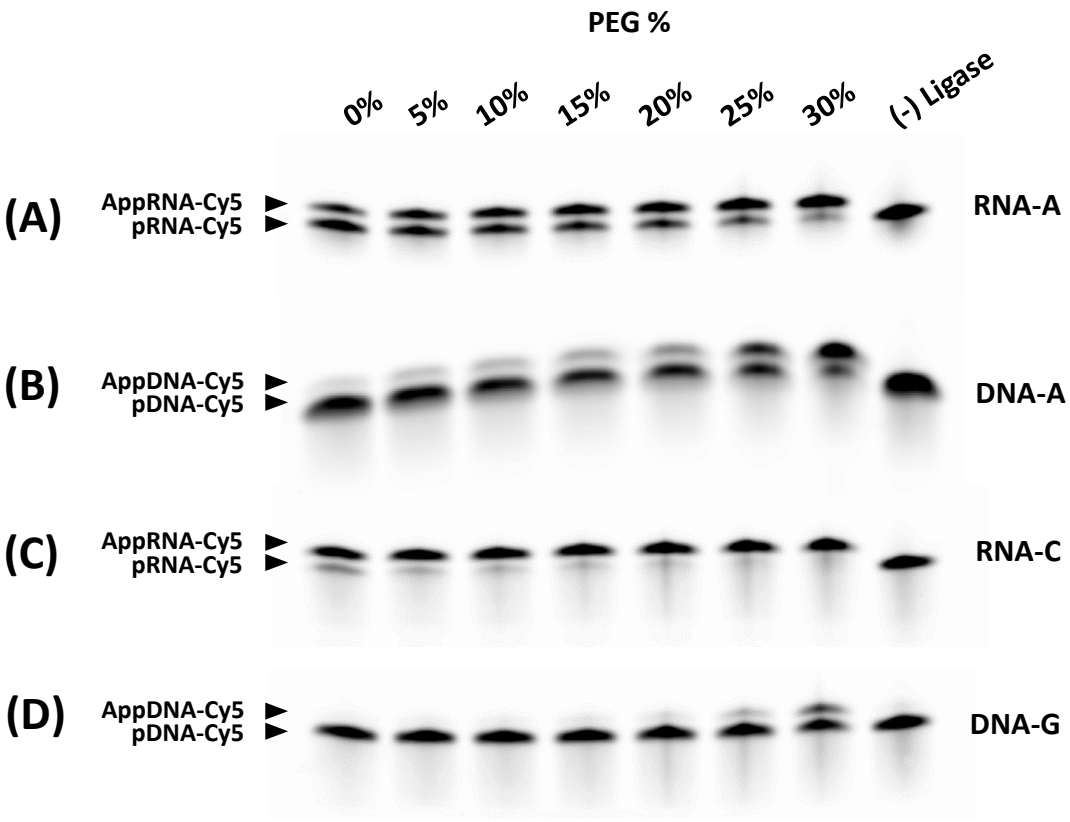

PEG %
0%
5%
10%
15%
20%
25%
30%
(-) Ligase
(A)
AppRNA-Cy5
pRNA-Cy5
RNA-A
(B)
AppDNA-Cy5
pDNA-Cy5
DNA-A
(C)
AppRNA-Cy5
pRNA-Cy5
RNA-C
(D)
AppDNA-Cy5
pDNA-Cy5
DNA-G

Figure S4.2. Adenylation of DNA-G and RNA-G adapters with extended incubation time.

Both DNA-G and RNA-G adapters are effectively adenylated within 2 hours.

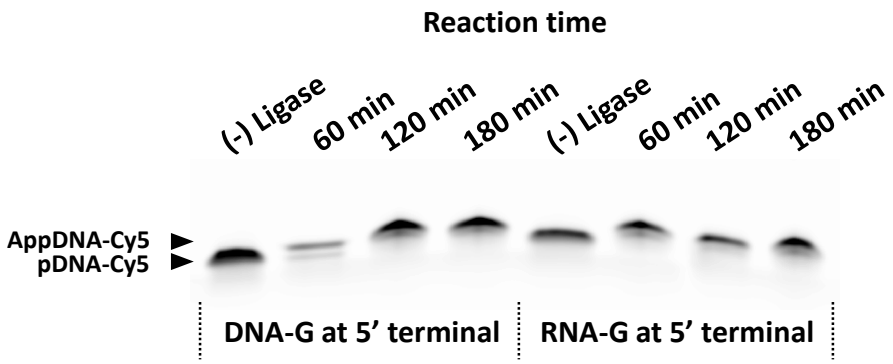

Reaction time
(-) Ligase
60 min
120 min
180 min
(-) Ligase
60 min
120 min
180 min
AppDNA-Cy5
pDNA-Cy5
DNA-G at 5' terminal
RNA-G at 5' terminal

Figure S4.3. Capture of RNA molecules within the total RNA sample. Total RNA extracted from pancreatic tissues is captured while being titrated from 1,000 ng to 10 ng (A). The three ligation products from the top, which represent RNA molecules captured by the adapter, are linearly proportional to the amount of input total RNA sample (B).

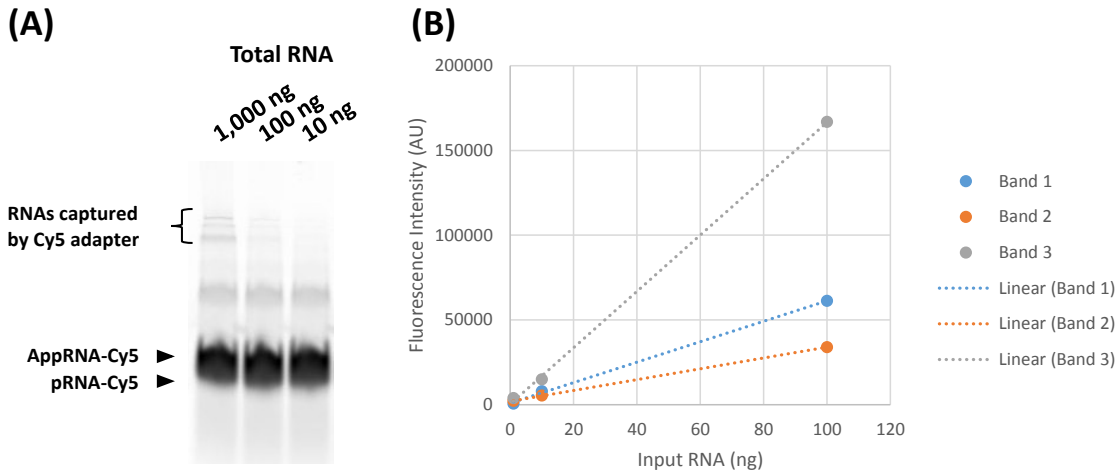

(A)
Total RNA
1,000 ng
100 ng
10 ng
RNAs captured
by Cy5 adapter
AppRNA-Cy5
pRNA-Cy5
(B)
Fluorescence Intensity (AU)
200000
150000
100000
50000
0
0
20
40
60
80
100
120
Input RNA (ng)
Band 1
Band 2
Band 3
Linear (Band 1)
Linear (Band 2)
Linear (Band 3)

Figure S4.4. Demonstration of microRNA – adapter ligation using pre-adenylated adapter synthesized with Archaeal RNA Ligase.

Both Cy5-spectra scan (excitation at 633 nm and emission at 670 nm) and Cy3-spectra scan (excitation at 532 nm and emission at 580 nm) are displayed. DNA-A adapter is effectively adenylated (lane 1: unadenylated adapter; lane 2: adenylated adapter) using the recommended method. Let-7a microRNA is effectively ligated to the adapter with and without addition of 100 ng of total RNA extracted from human pancreatic tissue (lane 4 and 6, respectively). Large RNA molecules within the total RNA are also captured by the adapter (lane 5). No de-adenylation of the pre-adenylated adapter is observed.

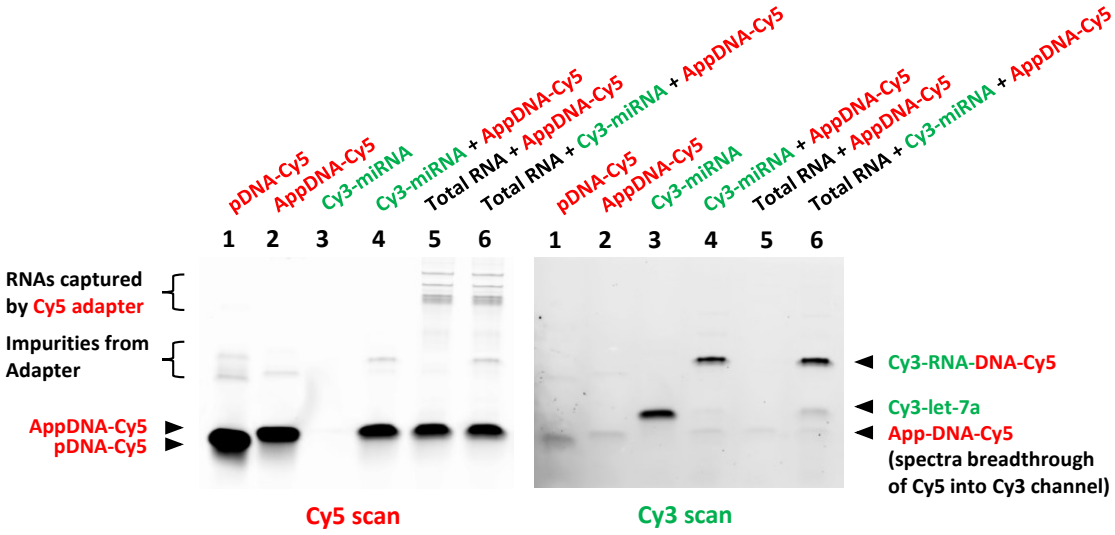
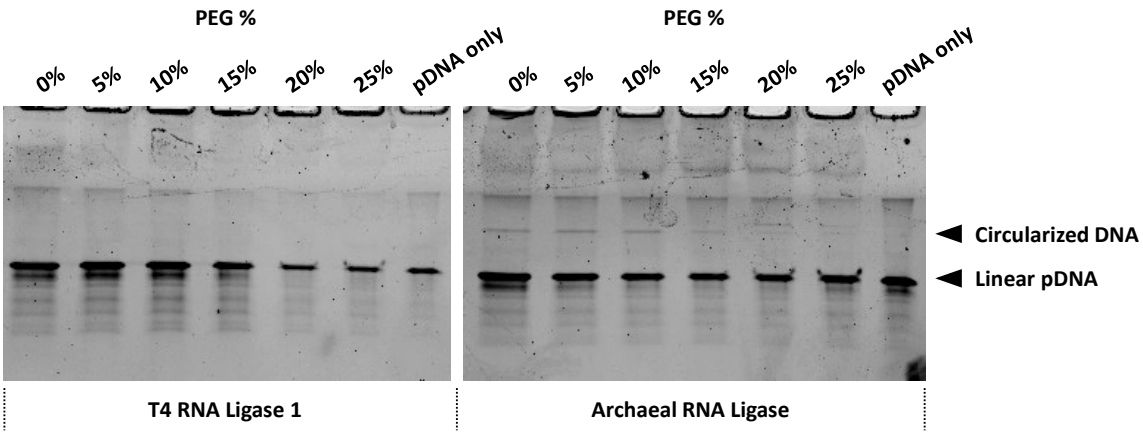

pDNA-Cy5
AppDNA-Cy5
Cy3-miRNA
Cy3-miRNA + AppDNA-Cy5
Total RNA + AppDNA-Cy5
Total RNA + Cy3-miRNA + AppDNA-Cy5
1 2 3 4 5 6
RNAs captured by Cy5 adapter
Impurities from Adapter
AppDNA-Cy5
pDNA-Cy5
Cy5 scan
Cy3 scan
Cy3-RNA-DNA-Cy5
Cy3-let-7a
App-DNA-Cy5
(spectra breadthrough of Cy5 into Cy3 channel)

Figure S4.5. Comparison of pre-adenylated adapter synthesized by T4 RNA Ligase 1 and Archaeal RNA Ligase in term of DNA circularization.

1 μM of single stranded DNA (80 nt) with 5'-phosphate and 3'-OH is mixed into the microRNA – adapter ligation reaction. The gel is stained with SYBR® Safe DNA Gel Stain. The DNA, but not let-7a and the adapter, is stained and visualized efficiently. Circularized DNA is generated when pre-adenylated adapter synthesized with Archaeal RNA Ligase is used in the reaction. The sample is excited with 488 nm laser and the fluorescence obtained through 526 nm emission filter.


PEG %
0%
5%
10%
15%
20%
25%
pDNA only
PEG %
0%
5%
10%
15%
20%
25%
pDNA only
Circularized DNA
Linear pDNA
T4 RNA Ligase 1
Archaeal RNA Ligase

Chapter 5: Capturing microRNA

Song,Y., Liu,K.J. and Wang,T.H. (2014) Elimination of ligation dependent artifacts in T4 RNA ligase to achieve high efficiency and low bias microRNA capture. PLoS One, 9, e94619.

5.1. Introduction

MicroRNA and other small RNA have added a new dimension to the connection between genotype and phenotype. These new mechanisms for gene expression regulation have led to a wealth of studies detailing the pervasive roles of microRNA in areas such as developmental biology,(95, 96, 97) stem cell biology (98, 99), cancer,(100, 101, 102) and plant genomics.(103, 104, 105) MicroRNA are studied both to elucidate their roles in fundamental mechanistic pathways as well as to develop novel disease biomarkers (106, 107) and therapeutics.(108) The majority of microRNA assay techniques been adapted from existing mRNA analysis methods. However, due to their short length, the first step of nearly all microRNA assays is to modify the microRNA through reverse-transcription,(109, 110) poly(A)-tailing (111, 112) or ligation.(89, 113) Among these methods, microRNA capture through adapter ligation is a pervasive first step in many PCR-,(107, 114) microarray-,(115, 116) bead- and sequencing- based assays.(113, 117, 118)

Due to the rising popularity of 2nd generation sequencing for small RNA detection and discovery, a number of studies have sought to benchmark

microRNA expression profiles across various detection platforms and systematically look for sources of bias.(113, 119, 120, 121, 122) Sequencing based methods are enjoying rising popularity due to their ability to identify small RNA species *de novo* and due to their ability to distinguish closely related isoforms. Although these sequencing approaches typically involve many sequential enzymatic steps including reverse transcription, PCR amplification, ligation, and poly(A) extension, a number of recent studies have pinpointed adapter ligation as the main contributor to expression profile bias.(85, 123, 124)

Ligation bias is critical because it underlies such a large number of microRNA analysis methods. Ligation can introduce two distinct levels of bias to microRNA expression profiles. First, bias can be introduced across samples when different adapters are used on different individual samples. Alon *et al* showed that consistent differential expression profiles can be seen across samples when the same adapter sequence is used but that large variations are seen when different adapter sequences are used even within the same sample. This can be a significant problem when comparisons are made across assay platforms that use different adapter sequences or when adapters are used to barcode individual samples such as in multiplexed deep sequencing applications. Second, bias can be introduced within each sample across various microRNA species, distorting the resultant expression profiles.

Hafner *et al* demonstrated that microRNA species can appear over or under-expressed by multiple orders of magnitude due to biases in ligation efficiency.(84, 124) This is less of an issue in differential expression analysis but is a significant issue when comparisons are made across microRNA species to rank expression levels. Though the majority of recent studies have examined bias in the context of sequencing based methods, this ligation bias will have similar effects on other microRNA assays such as PCR and array based methods that incorporate 3' ligation.

Recent studies have sought to identify the cause of ligation bias and remediate it.(84, 85, 123, 125) All of these recent studies have focused on improving adapter design to reduce bias. Jayaprakash *et al* found that two terminal bases on the 3' adapter can have dramatic effect on ligation efficiency.(85) Zhuang *et al* and Hafner *et al* demonstrated that secondary structure interactions can contribute significantly to variations in ligation efficiency.(84, 86) Remediation strategies have included optimization of ligase choice, optimizing secondary structure interactions, and incorporating adapter pools.(85, 86, 123, 125, 126) To our knowledge, few studies other than Zhang *et al* have investigated the optimization of reaction conditions for bias suppression.(126) Furthermore, nearly all have resorted to the use of randomized adapter pools.(85, 86, 123, 125, 126) This may be due to the

commonly held perception that T4 RNA ligase is inherently biased and difficult to use.

In this study, we have designed a microRNA capture method based on 3' adapter ligation that achieves very high efficiency (86% AVG) and low bias (10% SD) across all microRNA species tested. High efficiency capture is demonstrated even with microRNA that previous studies have had difficulty capturing and even based on the standard 3' modban adapter that previous studies have shown to exhibit high ligation bias. Using a panel of 20 microRNA, we studied key assay parameters such as PEG%, enzyme selection, adapter saturation, and design and show that they can be used to suppress bias and nearly eliminate ligation preference given suitable optimization methodology. We demonstrate that optimization must be done in the presence of total RNA using a microRNA panel to minimize global bias, as erroneous conditions can be found if optimization is done using only a single microRNA or in synthetic conditions.

5.2. Materials and Methods

Adapter Oligonucleotides. MicroRNA targets were synthesized by Integrated DNA Technologies (Coralville, IA). They consist of HPLC purified, DNA oligonucleotides with 5'-Ph and either 3'-ddC blocking group or 3'-Cy5 label. The adapters are based on modified 3' modban adapters.(88) The adapters were enzymatically pre-adenylated with T4 RNA ligase using a process

similar to Thomson *et al.*(127) Additional adapters were also synthesized for comparison purposes based on the SR1 and SR1-S sequences reported by.(86)

Synthetic microRNA. MicroRNA were synthesized by Integrated DNA Technologies (Coralville, IA). The sequences were taken from miRBase (www.mirbase.org) and are listed in **Table S5.1**. They targets consist of HPLC purified RNA oligonucleotides derivatized with 3'-OH and 5'-Cy3 end groups.

Ligation Protocol. Unless otherwise indicated, ligation was performed by mixing 1.25 μL of 2 μM adenylated adapter, 1 μL of T4 RNA Ligase buffer (New England Biolabs, Ipswich, MA), 5 μL of 50% PEG8000, 1 μL of synthetic target, 0.5 μL of total RNA, 1 μL of T4 RNA Ligase 2 truncated K227Q (New England Biolabs, Ipswich, MA) and water into a 20 μL reaction volume. The reaction was then incubated at 25 ºC for 4 hours and heat denatured at 65 ºC for 20 minutes in a thermal cycler. In the experiments where different ligases were investigated, T4 RNA Ligase 2 truncated, T4 RNA Ligase 2 truncated R55K K227Q, and Thermostable 5' App DNA/RNA Ligase were all obtained from New England Biolabs. In spiking experiments, 500 ng of human brain total RNA (Ambion, Austin, TX) was added to each sample.

PAGE Analysis. The samples were analyzed on precast 15% TBE-urea polyacrylamide gels (Bio-Rad, Hercules, CA). 5 μL of sample was mixed with 5 μL of loading buffer and heated for 5 minutes at 95 ºC. The sample was then loaded into the gel and run for either 30 min or 50 min at 300V. The separated gels were scanned using a Typhoon 9410 variable mode imager (GE Healthcare, Piscataway, NJ). The gel images were analyzed using ImageQuant (GE Healthcare, Piscataway, NJ) to obtain lane profiles. These profiles were then curve-fit with Gaussian curves using Origin (OriginLab, Northampton, MA) to precisely determine band position and intensity.

5.3. Results and Discussion

5.3.1. Bias in Ligation Based microRNA Capture

MicroRNA consist of short RNA sequences that are typically 17 - 23 nt in length. Due to their short length, 5' and/or 3' adapter ligation is often used to label, capture, or lengthen the microRNA before downstream detection. A number of studies have suggested a pooled adapter approach to average out the intrinsic effects of ligation bias. However, by improving ligation reaction design and optimization methodology, we have developed a ligation based

method that achieves high efficiency and low bias microRNA capture without the need for adapter pools.

As shown in **Figure 5.1**, an adapter oligonucleotide is ligated to the 3'-OH of each microRNA using T4 RNA ligase 2. To reduce side product formation, the adapter is first enzymatically pre-adenylated such that the ligation reaction can be performed in the absence of ATP. This prevents the microRNA in the sample from undergoing self-circularization, self-polymerization, and ligation to RNA species other than the adapter. Second, the 3' end of the adapter is blocked with dideoxycytosine (ddC), a fluorophore, or other moiety to prevent self-circularization and adapter concatenation. Finally, a recombinant mutant ligase is used. These enzymes lack the domain necessary for ATP incorporation and contain point mutations that further suppress side product formation. Such an approach is commonly used in microRNA analysis ahead of reverse transcription.(88, 89, 107, 128) Using this general reaction design, we investigated the effects of specific reaction conditions in suppressing ligation bias.

In order to characterize overall ligation efficiency and ligation bias, we synthesized a panel of 20 representative microRNA. Ten of the microRNA were selected based on their reported roles as important cancer-related microRNA (let-7a, miR-16, miR-21, miR-26a, miR-29b, miR-34a, miR-15a, miR-17p, miR-92a, and miR-155).(102) Seven of the microRNA were chosen

to enable comparison against recent publications (miR-31, miR-338, miR-567, miR-4803, miR-5183, miR-712, and miR-106b). For example, Zhuang *et al* reported difficulties capturing miR-4803, miR-5183, and miR-567 while Hafner *et al* reported low capture efficiencies for miR-31, miR-712, and miR-338.(84, 86) Jayaprakash *et al* reported that miR-106b could not be captured consistently under any of their experimental conditions. The final three microRNA were randomly selected (miR-25, miR-125b, miR-19b). Each target microRNA was labeled with a Cy3 dye at the 5' end to enable quantification of ligation efficiency under spiking conditions in total RNA. The reaction products were analyzed using denaturing PAGE, and the gels were scanned using a multimode imager. Image analysis was then used to obtain band positions and DNA quantity. Using this PAGE analysis method, we obtain excellent quantification and reproducibility. Quantification is linear from <5 amols to >10 pmols with an experiment to experiment CV of 10% (**Figure S5.1**).

5.3.2. Ligase Type

The first parameter we investigated was ligase type, as the ligases themselves likely have different intrinsic bias. **Figure 5.2** shows a denaturing PAGE analysis of adapter ligation on our 20 microRNA test panel using 4 different RNA ligases. For each ligase, we used the manufacturer recommended ligation conditions. As each microRNA is labeled with Cy3, the

capture efficiency was easily quantified using image analysis to compare the band intensities between the free microRNA band at ~20 nt and the ligated microRNA band at ~40 nt. A quantitative analysis of ligation efficiency for each microRNA and each ligase is provided in **Figure S5.2**.

T4 RNA ligase 2 truncated (T4 Rnl2 T) is a mutant enzyme that lacks the domain necessary for ATP incorporation, which should significantly reduce side product formation when used with pre-adenylated adapters in the absence of ATP.(129) As evidenced by the bright uniform bands in Figure 5.2A, this enzyme gave high ligation efficiency (66% AVG) and low bias (11% SD) with every microRNA species being captured at >40%. However, a significant number of background products can also be seen on the gel. A number of unique and randomly sized bands, which run faster than both the captured microRNA and the free adapter, are visible in each lane. It is unclear what these products are as typically only circularized products can run faster than linear products of the same size, but the same 5'-Cy3 that enables visualization should also prevent the formation of such circularized products.

T4 RNA ligase 2 truncated K227Q (T4 Rnl 2 TK) contains a point mutation that is designed to further reduce side product formation.(129, 130) This effect is clearly seen in Figure 5.2B, where side product formation is suppressed and only the desired products are visible. Smaller, randomly sized

bands are no longer apparent. Yet, the overall ligation efficiency has decreased dramatically (20% AVG) and ligation bias across the microRNA panel is quite significant (25% SD). Six microRNA were captured at <2% efficiency.T4 RNA ligase 2 truncated KQ (T4 Rnl 2 TKQ) is a double-point mutant that is also designed to have low side product formation but with increased ligation activity that is restored to the levels of T4 Rnl2 T.(129) In our experiments, little difference was seen between T4 Rnl 2 TK and T4 Rnl 2 TKQ, which had a capture efficiency of 17% ± 24%. T4 Rnl2 TKQ had 7 microRNA that were poorly captured at <2% efficiency.

Finally, we tried Thermostable 5' App DNA/RNA Ligase (MthRnl) from New England Biolabs, which is a point mutant of RNA ligase isolated from *Methanobacterium thermoautotrophicum*. This thermostable ligase is unable to incorporate ATP and works optimally at 65 ºC. Ligation at an elevated incubation temperature could serve to reduce bias from secondary structure interactions. Though the same adapters and microRNAs were used, a different pattern of ligation bias was seen. This likely arose from ligation preferences intrinsic to the ligase itself. In addition, despite the higher reaction temperature, neither ligation efficiency (30% AVG) nor bias (28% SD) was significantly improved over the T4 Rnl2 variants.

Although T4 Rnl2 T had high ligation efficiency and low bias, the large amounts of background products complicate downstream processes and were

deemed unacceptable. We chose to use T4 Rnl2 TK moving forward due to the belief that it would be easier to increase reaction efficiency and reduce bias than to suppress side product formation.

5.3.3. PEG Levels

Additives such as polyethylene glycol (PEG) (131, 132) and dimethyl sulfoxide (DMSO) (89, 115) are commonly added to ligation reactions to increase reaction efficiency. In our preliminary experiments, we saw minimal effect with DMSO addition (data not shown). However, we saw dramatic effects on ligation efficiency and bias due to PEG. PEG is thought to increase molecular crowding,(131, 132, 133) and many studies as well as manufacturer protocols have recommended ~15% PEG as an ideal concentration.(89, 107) Ligation efficiency is said to plateau or even decrease at high PEG levels.

Based on the initial results of Figure 5.2B, we optimized the effects of PEG on a subset of our 20 microRNA panel. We tested three microRNAs that were initially poorly captured by T4 Rnl2 TK, miR-31 (2.5% capture efficiency), miR-155 (28% capture efficiency), and miR-4803 (3.5% capture efficiency). 10 nM of each microRNA was individually spiked into 500 ng of total RNA. This represents about ~3-4 million microRNA copies per cell and is sufficient to approximate the aggregate expression of all microRNA within the cell. As the PEG concentration was varied from 0 - 35%, **Figure 5.3** and **Figure S5.3** show that the capture efficiency generally increased as a

function of PEG level and then decreased at high PEG levels. MiR-31 and miR-4803 behaved similarly, reaching the highest capture efficiency at 25% PEG, while miR-155 reached maximum efficiency at 15% PEG. The variation seen between these microRNA underscores the importance of optimizing reaction conditions across multiple microRNA species. The ideal conditions for a single microRNA do not necessarily extrapolate to other microRNA.

When the microRNAs were spiked into idealized buffer conditions rather than total RNA, a different behavior was seen. Maximum capture efficiency was reached at lower PEG levels with miR-31 and miR-4803 behaving similarly again, reaching plateau at 20% PEG, and miR-155 reaching plateau at 15% PEG. The overall ligation efficiencies also increased significantly, particularly for miR-155. Total RNA likely contains inhibitors that prevent ligation from reaching completion such as: 1) RNA species that bind and sequester the microRNA, 2) RNA species that ligate competitively to the microRNA, 3) RNA species that ligate competitively to the adapter, 4) RNA species that ligate competitively to each other, and 5) inhibitors of the ligase. Though our and reaction design should minimize effects 2), 3), and 4), the large amount of background RNA can still occupy the ligase binding site even if the actual ligation cannot occur (i.e. ligase shaking hands but not making deals). In addition, the decrease in ligation efficiency seen at high PEG levels under spiking conditions is also absent or greatly reduced under

idealized buffer conditions. It is unclear what the exact mechanism of PEG is and why different effects would be seen with and without total RNA spiking. Furthermore, the discrepancy in ligation behavior across the three microRNA also appears to decrease, with all three microRNA being well captured at 20% PEG.

This data illustrates that the optimal assay parameters determined using 1) a single microRNA vs. a panel and 2) in idealized buffer vs. spiking conditions are quite different, highlighting the critical importance of optimization methodology and design. Most previous publications, as well as manufacturer protocols, have recommended 12-15% PEG. At this PEG level, only miR-155 is optimally captured under spiking conditions. In spiking conditions, ligation at 20-25% PEG is optimal whereas ligation at the recommended 15% PEG leads to very low efficiency. Of all the parameters investigated, PEG had the most dramatic effect on ligation bias and efficiency.

5.3.4. Adapter Concentration and Ligase Amount

Next we tested the effects of adapter concentration and ligase amount in conjunction with one another. Adapter concentration must be in excess to the ligated species to drive ligation forward. In addition to microRNA, samples often contain other RNA species such as mRNA, rRNA, and siRNA that can also be ligated. Having too few adapters will limit the ligation efficiency and increase bias. Having too large an excess of adapters will promote side

product formation. In addition, large excesses of free adapters may also complicate downstream assay processes. The ligase amount also needs to be sufficient to obtain a high ligation efficiency in a reasonable amount of time. However, high ligase amounts can promote side product formation as well. Practically speaking, ligase is the most expensive reaction component and should be minimized to reduce costs.

Adapter concentration and ligase amount were optimized by quantifying their effects on the capture of miR-31, both in presence and absence of total RNA. **Figure 5.4** and **Figure S5.4** show the effects of concurrently varying adapter concentration from 50 nM to 400 nM and ligase amount from 100 units to 400 units. In the absence of total RNA, little effect was seen by increasing either adapter or enzyme levels. Even at 50 nM adapter and 100 units of enzyme, ligation efficiency was 88% due to their high excess. However, under the same conditions in the presence of total RNA, ligation efficiency drops to 34% due to the large amounts of other RNA species in solution that the adapters can be ligated to and that ligase can spend time shaking hands with. Under these realistic spiking conditions, much larger amounts of adapter and enzyme were needed; 200-300 nM of adapter and 200 units of enzyme were needed to reach saturation. Although the highest capture efficiency was obtained with high amounts of both

adapter and enzyme, high levels of at least one component also gave relatively high efficiencies.

5.3.5. Incubation Time

For the previous reactions, a 4 hour incubation was performed. We investigated whether this time was sufficient and whether it could be reduced. Using miR-31 as a model and the optimized protocol developed thus far, we tested incubation times from 30 minutes to 18 hours both in the presence and absence of total RNA. **Figure 5.5** and **Figure S5.5** illustrate the microRNA capture efficiency as a function of time. As expected, the ligation reaction proceeded much faster in the absence of total RNA, reaching >80% in only 30 minutes and reaching plateau in <2 hours. When spiked into total RNA, plateau was not reached until >8 hours. This was likely due to the greater number of RNA species within the sample that the adapters could be ligated to. By 4 hours, ligation reached 92% of the plateau value, striking a good compromise between incubation time and ligation efficiency.

5.3.6. Incubation Temperature

Secondary structure has been proposed as a main contributor to ligation bias.(84, 86) We investigated the effect of incubation temperature on ligation efficiency and ligation bias under the premise that elevated incubation temperatures could potentially reduce secondary structure interactions and alleviate bias. Initially, when we performed ligation experiments at 65 °C

using MthRnl, we did not see any significant difference when compared to ligation at 25 °C using the T4 Rnl2 variants. However, in this case both the temperature and the ligase were changed. The intrinsic bias of the MthRnl could have swamped out any effects due to temperature.

We performed a second analysis using T4 Rnl2 TK while incubating at 4 °C for 18 hours, 25 °C for 4 hours, or 37 °C for 4 hours. T4 Rnl2 TK is not thermostable and is denatured at 65 °C so temperatures beyond 37 °C were not tried. Incubation at 4 °C allows for the reaction to proceed for an extended amount of time to compensate for decreased enzyme activity. Most enzymes recommend incubation at 25 °C. We also performed ligation at 37 °C to see if a modest increase in incubation temperature would have any effect on bias. **Figure 5.6** and **Figure S5.6** show a graph and gel images, respectively, of the ligation efficiency for all 20 microRNA in our test panel using the 4 °C, 25 °C, and 37 °C conditions when spiked into 500 ng of total RNA.

The highest efficiency and lowest bias were seen at the 25 °C condition where ligation efficiency across the 20 microRNA panel was 53.1% AVG ± 13.7% SD. When the incubation temperature was decreased to 4 °C, the ligation efficiency across the panel decreased to 23.6% AVG ± 7.2% SD. Despite the increased ligation time (18 hours vs. 4 hours), the decreased enzyme activity at 4 °C led to significantly reduced ligation efficiency.

Interestingly, the capture efficiency CV for the 4 ºC condition (31% CV) was only slightly worse than at 25 °C (26% CV), indicating that temperature does not have a large impact on bias across this range. The pattern in capture efficiency for each individual microRNA in the panel was fairly similar for the two conditions except that the capture efficiency at 25 °C was 2x higher in most cases. When the incubation temperature was increased to 37 °C, the ligation efficiency across the panel dropped to 34.7% AVG ± 17.7% SD. The higher incubation temperature reduced the ligase activity but unexpectedly increased the bias over both the 4 °C and 25 °C conditions to nearly 51% CV. Incubation temperature likely results in a combination of effects on ligase activity, ligase degradation, and secondary structure formation that impact ligation efficiency in a complex manor. With a few exceptions, the pattern in ligation efficiency across the panel was generally similar to the 4 ºC and 25 ºC conditions.

5.3.7. Adapter Design

Thus far, we've investigated the use of reaction conditions to suppress the intrinsic bias of T4 Rnl2 TK. Adapter design can also play a large part in ligation bias due to primary sequence and/or secondary structure effects. However, the differences in microRNA capture efficiency seen across published studies illustrate the unpredictable nature of these interactions. For example, even when adapters were logically designed to either eliminate

inhibitory secondary structures or promote favorable interactions, only modest improvements, if any, were seen.(86) In **Figure 5.7** and **Figure S5.7**, we performed adapter ligation on the 20 microRNA panel in spiking conditions (500 ng total RNA) using four different adapter sequences. First, we synthesized two versions of our modified modban adapter (88) to test whether having RNA or DNA as the 5' residue would affect ligation efficiency or bias. As the ligases used herein are RNA ligases used for single stranded blunt end ligations, it is possible that the ligases will have a preference for ligating RNA versus DNA. The rA version contains a ribo-A as the 5' base while the dA version contains a dexoyribo-A as the 5' base. As seen in Figure 5.7 left, both of our modified modban adapters achieved high ligation efficiencies with low bias. No significant difference was seen between the rA or dA versions of the adapter, indicating that T4 Rnl2 TK displays no real preference for RNA-RNA ligation or RNA-DNA ligation. The rA adapter had a capture efficiency of 68.5% ± 17.6% (AVG ± SD) across the 20 microRNA panel while the dA adapter had a capture efficiency of 71.6% ± 15.4% (AVG ± SD). Even under spiking conditions, the capture efficiencies surpass what previous publications were able to achieve using complex adapter pool strategies in idealized buffer conditions. Using the dA adapter, 19 microRNA were captured at >50% with the lowest still being captured at 31%. We attribute the low capture efficiency of miR-155 to degradation or synthesis problems. We re-synthesized miR-155 multiple times. Immediately after

receiving the target, high ligation efficiencies were obtained which would slowly degrade over time. MiR-155 was the only target we saw this effect with.

Encouraged by the previous results, we synthesized the SR1 and SR1-S adapters used by Zhuang (86) and Hafner (84) to test whether our reaction would work equally well with other adapter designs. SR1 is a standard adapter commonly used in Illumina's sequencing products. SR1-S was designed by Zhuang to reduce ligation bias by eliminating inhibitory secondary structure interactions. As seen in Figure 5.7 right, the overall ligation efficiency and bias were much poorer with both of these adapter designs. The SR1 capture efficiency across the 20 microRNA panel was 49.4% ± 20.6% with only 7 microRNA captured at > 50%. The SR1-S adapter fared even worse with only 1 microRNA being captured at >50%. SR1-S had a capture efficiency of 16.9% ± 15.2% (AVG ± SD) across the panel. This result parallels that reported by Zhuang where the SR1-S adapter unexpectedly performed much worse than the SR1 adapter and failed to improve ligation efficiencies despite eliminating secondary structure interactions. Given the current reaction conditions and microRNA test panel, our modified modban adapter appears to work significantly better than the SR1 adapter. It is possible that the ligation reaction conditions can be optimized specifically for the SR1 adapter but we did not attempt this.

5.3.8. High Efficiency and Low Bias microRNA Capture

Based on the optimized conditions determined in previous experiments, we performed ligation on the 20 microRNA panel using 300 units of T4 Rnl 2 TK, 25% PEG, 200 nM adapter incubated for 4 h at 25 °C. Ligation was performed both in idealized buffer and in total RNA spiking conditions as shown in **Figure 5.8**. Each experiment was performed in triplicate. In the absence of total RNA, the capture efficiency across the 20 microRNA panel was 86% AVG ± 10% SD. No microRNA was captured at less than 60% with 18 of 20 microRNA being captured at 70% or greater. This is a considerable improvement over the initial conditions of 20% ± 25% (AVG ± SD). Optimization of various parameters such as PEG, adapter, enzyme levels, and incubation time led to an increase in overall ligation efficiency and the concurrent benefit of effectively reducing ligation bias. Saturating levels of each parameter tend to push all of the competing reactions toward completion, regardless of moderate differences in thermodynamic equilibrium and kinetics. Thus, as the capture efficiency of all the microRNA increases and approaches 100%, bias is concurrently reduced since capture saturates at 100%.

When compared to previous studies performed in idealized buffer background, a significant improvement in average capture efficiency and bias was seen. For example Hafner *et al* reported that miR-16 (92%), miR-21(66%),

miR-155 (90%), and miR-567 (73%) were well captured while miR-31 (48%) and miR-338 (3%) were less well captured.(84) In our process, all of these microRNA are uniformly captured at 87%, 91%, 60%, 74%, 74%, and 93%, respectively. Similar results are seen when compared to Zhuang *et al*.(86) They reported let-7a (100%), miR-31 (87%), and miR-567 (56%) to be well captured while miR-4803 (1%), miR-5183 (42%), and miR-712 (8%) were poorly captured. All of these microRNA species are well captured with our methods at 97%, 74%, 74%, 90%, 97% and 69%, respectively. In addition, miR-106, which could not be captured consistently by Jayaprakash *et al* under any conditions, is now captured at 90%.(85)

These differences are further pronounced when comparing microRNA capture in realistic spiking conditions. The presence of total RNA likely reduces capture efficiency due to the aforementioned competitive reactions. Under realistic spiking conditions in presence of total RNA, our capture efficiency is slightly decreased and bias is slightly increased with a capture efficiency of 64% AVG ± 17% SD across the panel. 16 of 20 microRNA have capture efficiencies greater than 50%. Only miR-155 had a low capture efficiency (27%). We experienced repeatability problems with miR-155, likely due to synthesis and degradation, as previously described. In comparison, Zhuang *et el* reported significantly reduced capture efficiencies under spiking conditions with efficiencies of 19%, 17%, 9%, 2% and 4% for let-7a, miR-31,

miR-567, miR-4803, and miR-5183, respectively.(86) To resolve this, they proposed a random adapter pool approach to average out the low capture efficiencies. However, the capture efficiencies for the pooled approach still remained relatively low at 83%, 27%, 8%, 9%, and 14%, respectively. A recent study by Zhang *et al*(126) has also used randomized adapter pools to reduce bias. Bias was reduced to1.8-fold from the expected frequency across a 29 microRNA panel. In comparison, the current method uses only a single adapter and achieves highly efficient and low bias capture levels of 74%, 63%, 61%, 73%, and 83% for let-7a, miR-31, miR-567, miR-4803, and miR-5183. In fact, our spiking capture efficiencies surpass what has been demonstrated by previous studies under idealized buffer conditions.

The experiments shown in Figure 5.8 were performed using a dA adapter labeled with Cy5. Additionally, we also performed this experiment in triplicate using an rA adapter labeled with Cy5 and an rA adapter blocked with ddC (data not shown). In each case, similar results were obtained, demonstrating that the adapter ligation process is highly repeatable and robust. This is further evidenced by the low average experiment-to-experiment SD for each microRNA seen in Figure 5.8, which was 3% in idealized buffer and 7% in total RNA spiking. microRNA capture efficiency is also consistent across a broad range of microRNA input levels as seen in **Figure S5.8**.

5.4. Figures

Figure 5.1. Schematic illustration of microRNA capture by 3' adapter ligation.

The 19 nt, enzymatically pre-adenlyated adapter probe is ligated to the 3' OH of microRNA using T4 RNA ligase 2. The reaction is run at 25 °C for 4 hours in the absence of ATP. In order to characterize capture efficiency, the microRNA is end labelled with Cy3. The 3’ end of the adapter probe is blocked by –ddC, a fluorophore, or other moiety to prevent the formation of concatemers and circularized products.

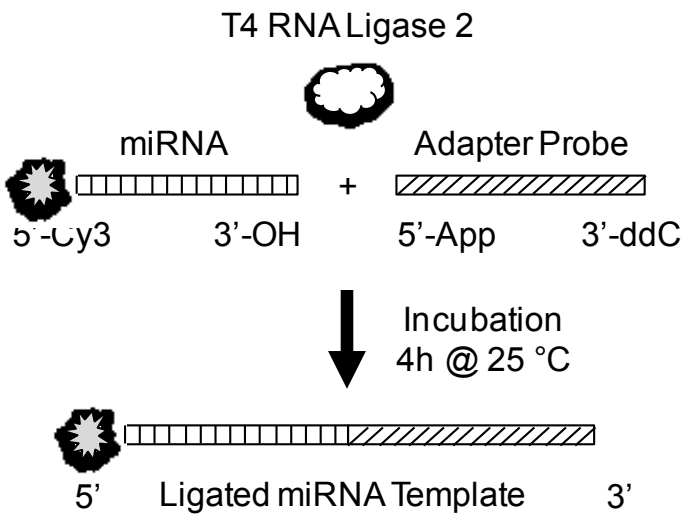

T4 RNA Ligase 2
miRNA
Adapter Probe
+
5'-Cy3
3'-OH
5'-App
3'-ddC
Incubation
4h @ 25 °C
5'
Ligated miRNA Template
3'

Figure 5.2. MicroRNA capture was performed with 4 different ligases using the vendor recommended protocols to compare capture efficiency across 20 different microRNA.

The ligation products were analyzed by 15% denaturing urea-PAGE. Capture efficiency was determined by performing a Cy3 scan and comparing the intensities of the ~40 nt captured microRNA band versus the ~20 nt free microRNA band. T4 RNA Ligase 2 truncated (T4 Rnl2 T) had high average capture efficiency and low bias but many randomly sized background products. The point mutant enzymes T4 RNA Ligase 2 truncated K227Q (T4 Rnl2 TK) and T4 RNA Ligase 2 truncated KQ (T4 Rnl2 TKQ) had decreased side product formation but also lower average capture efficiency and higher bias. Thermostable 5´ App DNA/RNA Ligase (Mth Rnl), which was performed at 65 °C instead of 25 ºC, had similar average capture efficiency and bias but with distinct ligation efficiency pattern.

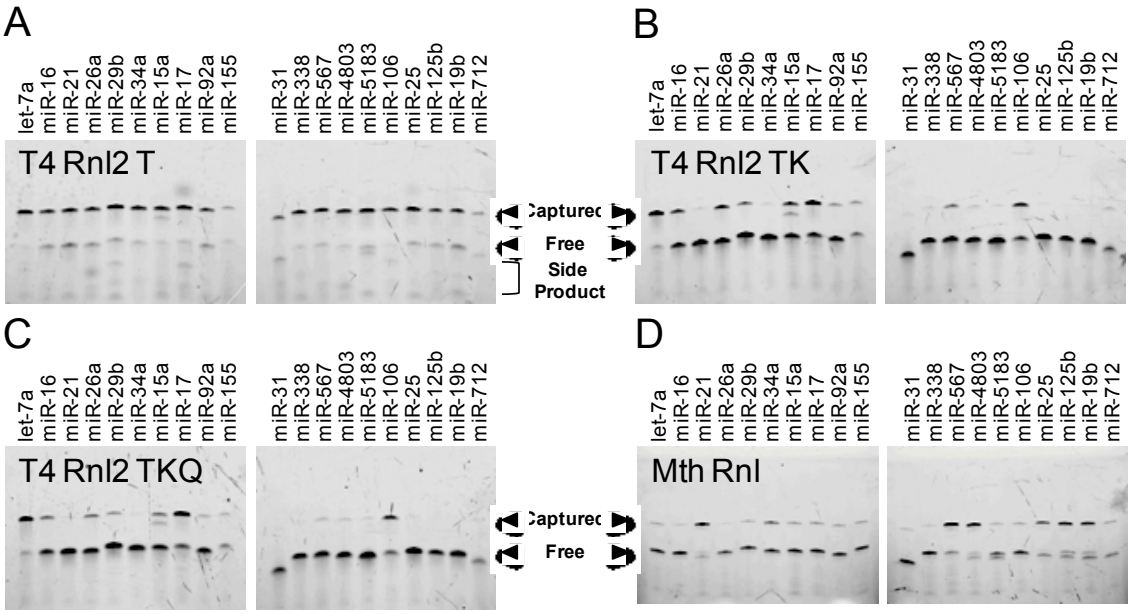

A
let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106
miR-25
miR-125b
miR-19b
miR-712
T4 Rnl2 T
Captured
Free
Side
Product
B
T4 Rnl2 TK
C
T4 Rnl2 TKQ
D
Mth Rnl

Figure 5.3. Comparison of ligation efficiency as a function of PEG percentage for miR-31, miR-155, and miR-4803 in idealized buffer (open markers) and total RNA spiking conditions (filled markers). miR-155 and miR-4803 display similar behaviour with respect to PEG percentage while miR-31 behaves distinctly. Different behaviour is also seen between idealized buffer conditions and total RNA spiking conditions, illustrating the importance of optimization methodology in extrapolating assay performance. Optimizations performed using a single microRNA species in idealized buffer may not extrapolate to other microRNA under actual assay conditions.

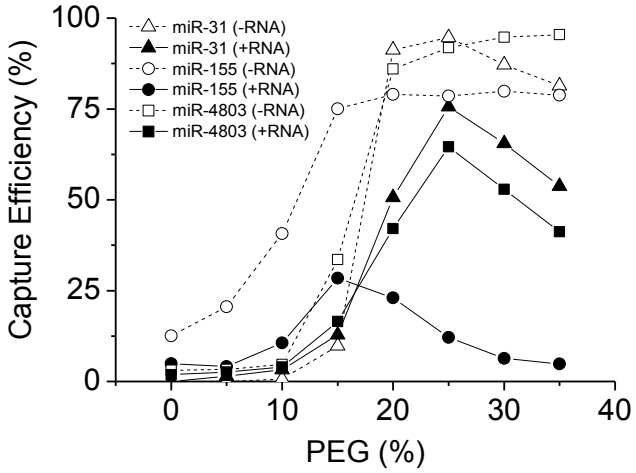

Capture Efficiency (%)
PEG (%)
miR-31 (-RNA)
miR-31 (+RNA)
miR-155 (-RNA)
miR-155 (+RNA)
miR-4803 (-RNA)
miR-4803 (+RNA)
0
25
50
75
100
0
10
20
30
40

Figure 5.4. Optimization of adapter probe and enzyme concentration

The adapter probe concentration and T4 Rnl2 TK amount were changed simultaneously to see their joint effect on the capture efficiency of miR-31. The experiment was performed under idealized buffer conditions and total RNA spiking conditions which lead to distinct conditions for optimum capture efficiency. Under spiking conditions, greater amounts of probe and enzyme are necessary to obtain high capture efficiency.

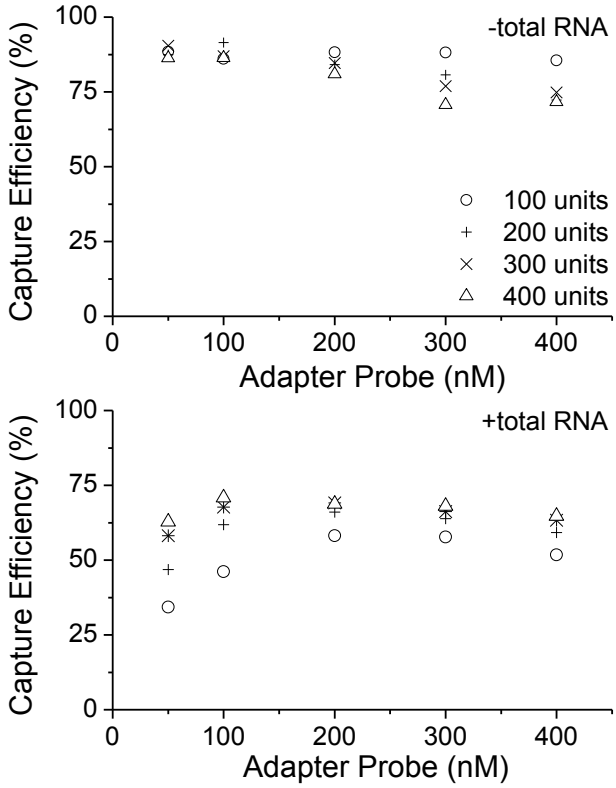

-total RNA
100 units
200 units
300 units
400 units
Capture Efficiency (%)
Adapter Probe (nM)
+total RNA
Capture Efficiency (%)
Adapter Probe (nM)

Figure 5.5. Reaction time optimization.

The ligation reaction was incubated at 25 °C for 30 min to 18 hours to investigate the effect of time on the capture efficiency of miR-31 in idealized buffer conditions and total RNA spiking conditions. In idealized buffer conditions, the ligation reaches completion in < 2 hours. Whereas in spiking conditions, the ligation does not reach full completion until >8 hours.

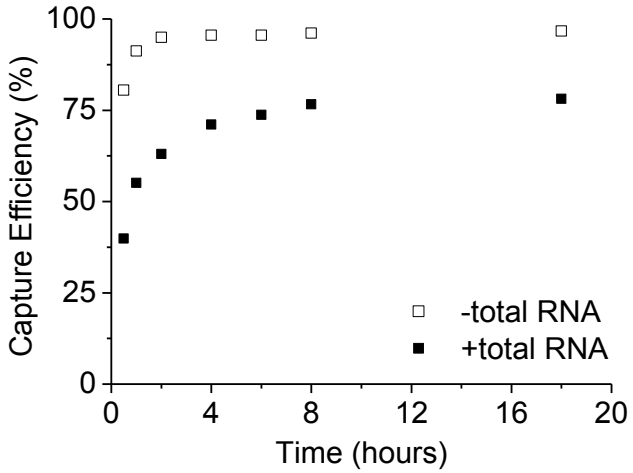

Capture Efficiency (%)
100
75
50
25
0
0
4
8
12
16
20
Time (hours)
-total RNA
+total RNA

Figure 5.6. Temperature optimization.

Comparison of incubation temperature on capture efficiency across the 20 microRNA panel. The ligation reaction was incubated for 18 hours at 4 °C, 4 hours at 25 °C or 4 hours at 37 °C under total RNA spiking conditions.

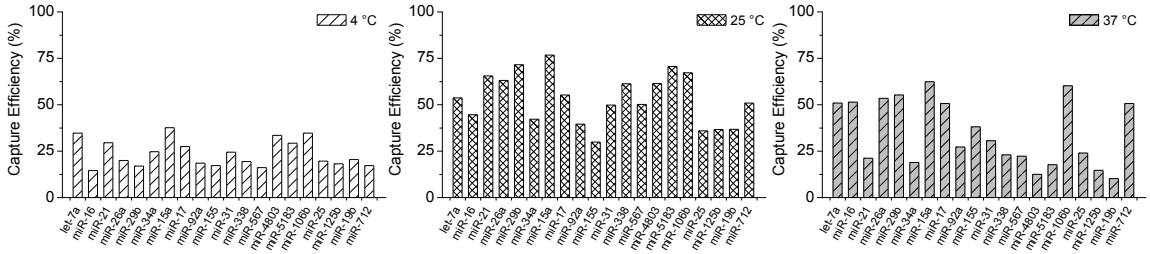

Capture Efficiency (%)
100
75
50
25
0
4 °C
let-7a
miR-16
miR-21
miR-26a
miR-26b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5163
miR-106b
miR-25
miR-125b
miR-19b
miR-712
25 °C
37 °C

Figure 5.7. Comparison of capture efficiency across the 20 microRNA panel spiked into 500 ng of total RNA using 4 different adapter probe designs.

The rA and dA adapter probes have identical sequences based on modified modban probe design except the 5' base is either RNA or DNA as indicated. The SR1 and SR1-S adapters are taken from Zhuang *et al* (39). T4 Rnl2 TK shows no preference for DNA or RNA at the ligation site. However, overall capture efficiency and bias were significantly worse for the SR1 and SR1-S adapters.

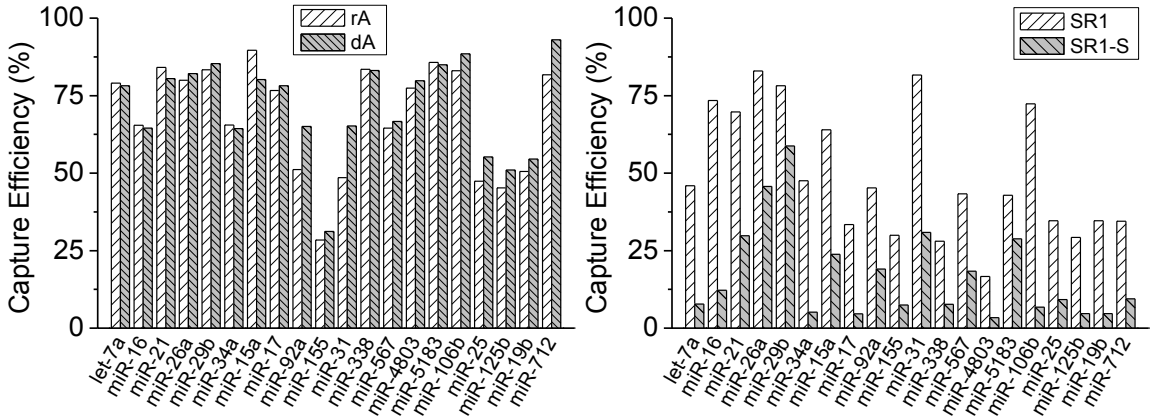
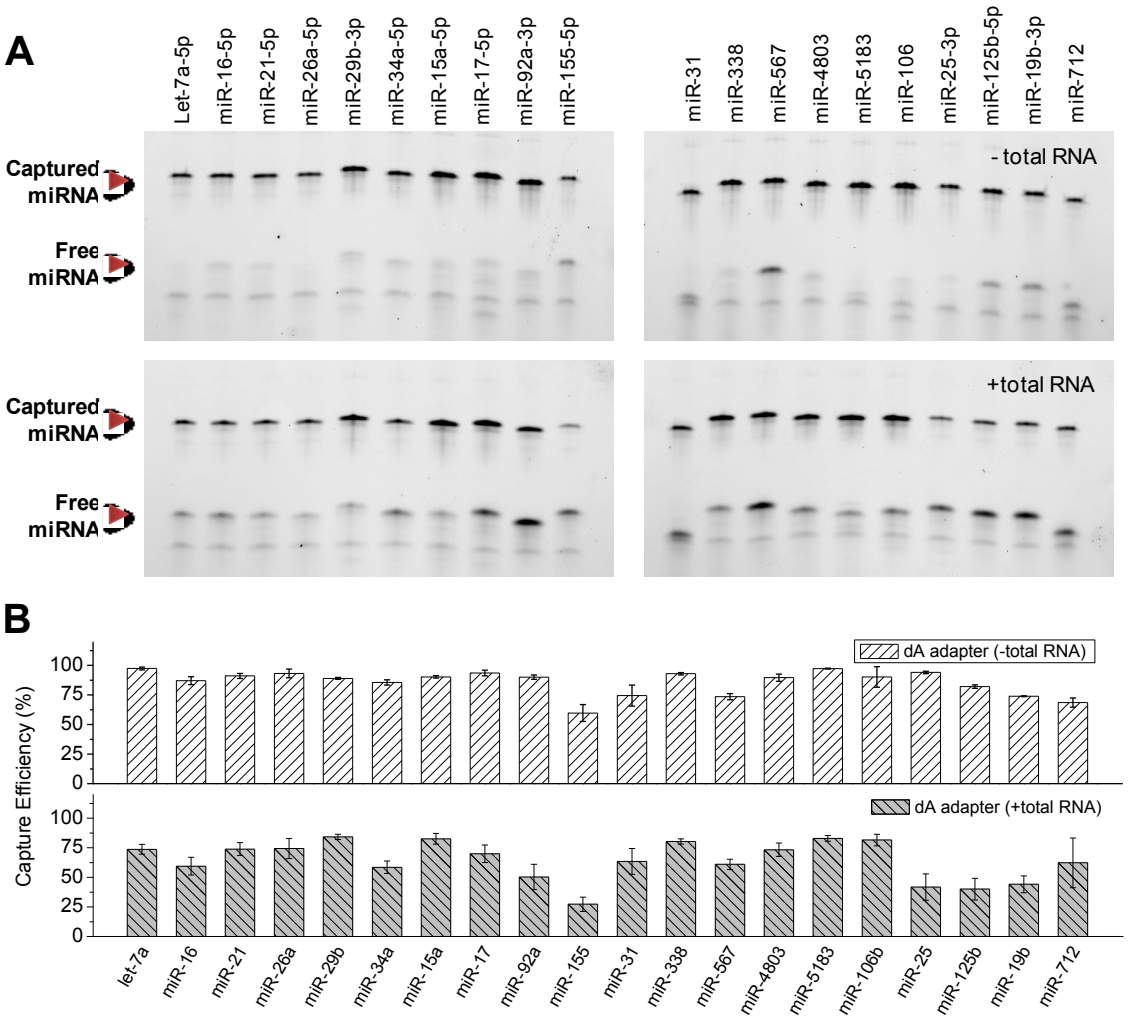

Capture Efficiency (%)
100
75
50
25
0
rA
dA
let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106b
miR-25
miR-125b
miR-19b
miR-712
Capture Efficiency (%)
100
75
50
25
0
SR1
SR1-S
let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106b
miR-25
miR-125b
miR-19b
miR-712

Figure 5.8. The optimized capture efficiency for the 20 microRNA panel in idealized buffer conditions and total RNA spiking conditions.

The data is presented as representative gel images and as graphs based on image analysis of 3 sets of independent experiments. High capture efficiency and low bias are obtained across the panel both in idealized buffer conditions and total RNA spiking conditions.


A
Let-7a-5p
miR-16-5p
miR-21-5p
miR-26a-5p
miR-29b-3p
miR-34a-5p
miR-15a-5p
miR-17-5p
miR-92a-3p
miR-155-5p
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106
miR-25-3p
miR-125b-5p
miR-19b-3p
miR-712
Captured miRNA
Free miRNA
- total RNA
+total RNA
B
Capture Efficiency (%)
dA adapter (-total RNA)
dA adapter (+total RNA)
let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106b
miR-25
miR-125b
miR-19b
miR-712

5.5. Supplementary Information

Supplementary table and figures follow below.

Table S5.1. Oligonucleotides

The adapter probe was based on a modified version of the modban adapter

Name	**Sequence 5'-> 3'**
let-7a	5'- /5Cy3/rUrG rArGrG rUrArG rUrArG rGrUrU rGrUrA rUrArG rUrU -3'
miR-106b	5'- /5Cy3/rUrA rArArG rUrGrC rUrGrA rCrArG rUrGrC rArGrA rU -3'
miR-125b	5'- /5Cy3/rUrC rCrCrU rGrArG rArCrC rCrUrA rArCrU rUrGrU rGrA -3'
miR-155	5'- /5Cy3/rUrU rArArU rGrCrU rArArU rCrGrU rGrArU rArGrG rGrGrU -3'
miR-15a	5'- /5Cy3/rUrA rGrCrA rGrCrA rCrArU rArArU rGrGrU rUrUrG rUrG -3'
miR-16	5'- /5Cy3/rUrA rGrCrA rGrCrA rCrGrU rArArA rUrArU rUrGrG rCrG -3'
miR-17	5'- /5Cy3/rCrA rArArG rUrGrC rUrUrA rCrArG rUrGrC rArGrG rUrArG -3'
miR-19b	5'- /5Cy3/rUrG rUrGrC rArArA rUrCrC rArUrG rCrArA rArArC rUrGrA -3'
miR-21	5'- /5Cy3/rUrA rGrCrU rUrArU rCrArG rArCrU rGrArU rGrUrU rGrA -3'
miR-25	5'- /5Cy3/rCrA rUrUrG rCrArC rUrUrG rUrCrU rCrGrG rUrCrU rGrA -3'
miR-26a	5'- /5Cy3/rUrU rCrArA rGrUrA rArUrC rCrArG rGrArU rArGrG rCrU -3'
miR-29b	5'- /5Cy3/rUrA rGrCrA rCrCrA rUrUrU rGrArA rArUrC rArGrU rGrUrU -3'
miR-31	5'- /5Cy3/rArG rGrCrA rArGrA rUrGrC rUrGrG rCrArU rArGrC rU -3'
miR-338	5'- /5Cy3/rUrC rCrArG rCrArU rCrArG rUrGrA rUrUrU rUrGrU rUrG -3'
miR-34a	5'- /5Cy3/rUrG rGrCrA rGrUrG rUrCrU rUrArG rCrUrG rGrUrU rGrU -3'
miR-4803	5'- /5Cy3/rUrA rArCrA rUrArA rUrArG rUrGrU rGrGrA rUrUrG rA -3'
miR-5183	5'- /5Cy3/rUrA rUrUrU rGrGrA rCrArA rArUrU rUrGrA rGrUrC rA -3'
miR-567	5'- /5Cy3/rArG rUrArU rGrUrU rCrUrU rCrCrA rGrGrA rCrArG rArArC -3'
miR-712	5'- /5Cy3/rCrU rCrCrU rUrCrA rCrCrC rGrGrG rCrGrG rUrArC rC -3'
miR-92a	5'- /5Cy3/rUrA rUrUrG rCrArC rUrUrG rUrCrC rCrGrG rCrCrU rGrU -3'
Modban Adapter*	5'- CTGTAGGCACCATCAAT -3'
SR1 Adapter	5'- TCGTATGCCGTCTTCTGCTTG-NH_2 -3'
SR1-S Adapter	5'- TCGTATGCCGTCACGGCTCGA-NH_2 -3'

Figure S5.1. Characterization of the gel scanner.

The 19 nt Cy5-labelled adapter probe was serially diluted and analyzed by 15% denaturing PAGE. A Cy5 scan was performed on a Typhoon imager, and image analysis was used to determine band intensity. The experiment was repeated 3 times. The fluorescence was linear from <5e-18 mol to >1e-11 mol. The weighted least squares fit shown in red has an R^2 ~0.999.

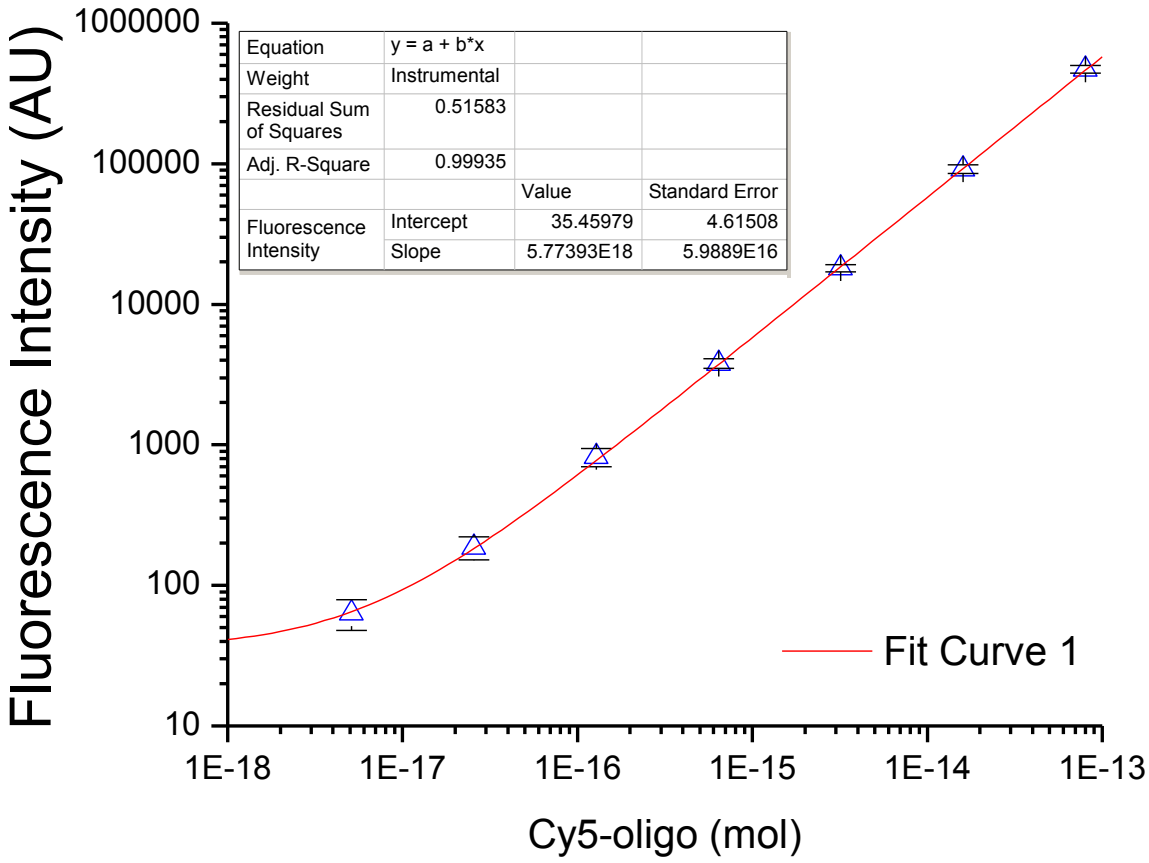

Fluorescence Intensity (AU)
1000000
100000
10000
1000
100
10
1E-18
1E-17
1E-16
1E-15
1E-14
1E-13
Cy5-oligo (mol)
Fit Curve 1
Equation y = a + b*x
Weight Instrumental
Residual Sum of Squares 0.51583
Adj. R-Square 0.99935
Value Standard Error
Fluorescence Intensity Intercept 35.45979 4.61508
Slope 5.77393E18 5.9889E16

Figure S5.2. Comparison of ligases in microRNA capture performance

Image analysis was performed on the gel data in Figure 5.2 to obtain capture efficiencies for each of the 20 microRNA in the panel using T4 Rnl2 T, T4 Rnl2 TK, T4 Rnl2 TKQ, and Mth Rnl. The reactions were performed using the vendor recommended protocols in each case.

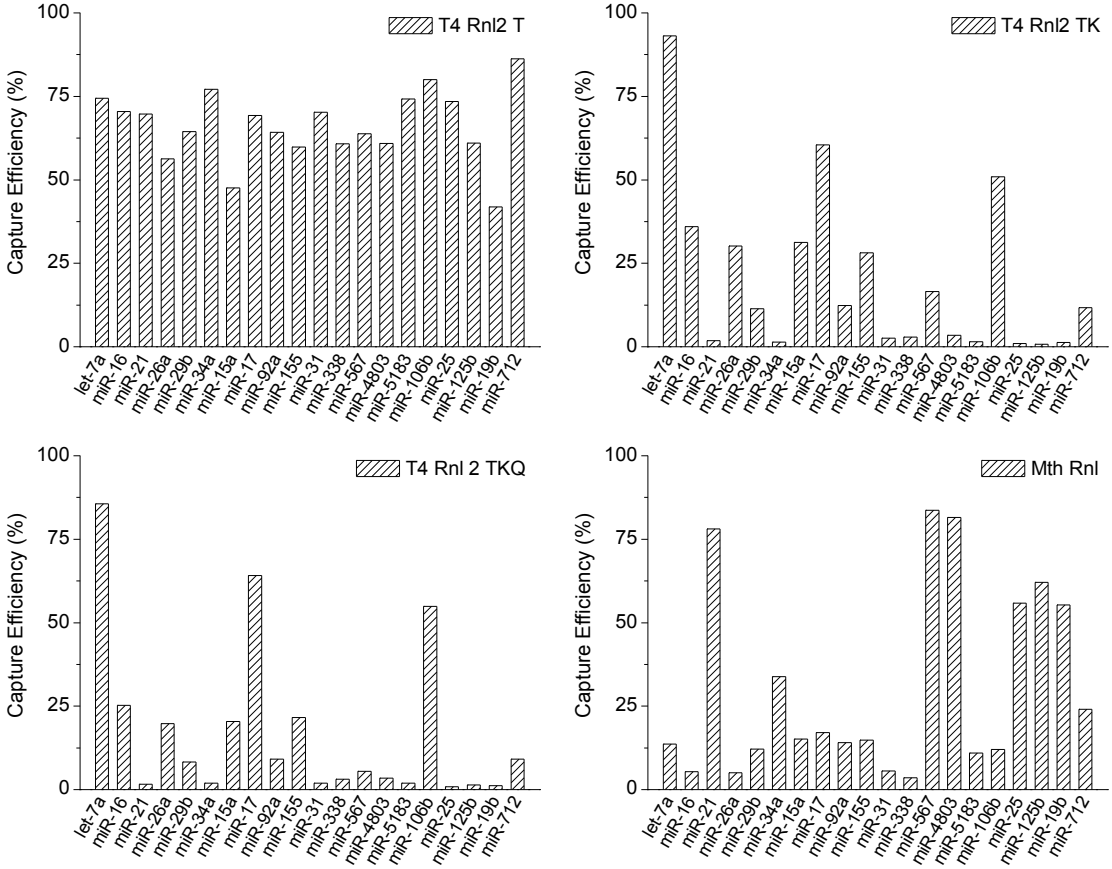

T4 Rnl2 T
T4 Rnl2 TK
T4 Rnl 2 TKQ
Mth Rnl
Capture Efficiency (%)
0
25
50
75
100
let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106b
miR-25
miR-125b
miR-19b
miR-712

Figure S5.3. Optimization of PEG percentage

Raw PAGE images depicting the effects of PEG on capture efficiency for *miR-31*, *miR-155* and *miR-4803* in idealized buffer conditions (-RNA) and spiked into 500 ng of total RNA (+RNA). PEG was varied from 0% to 35%. Image analysis was performed on this data to obtain Figure 5.3.

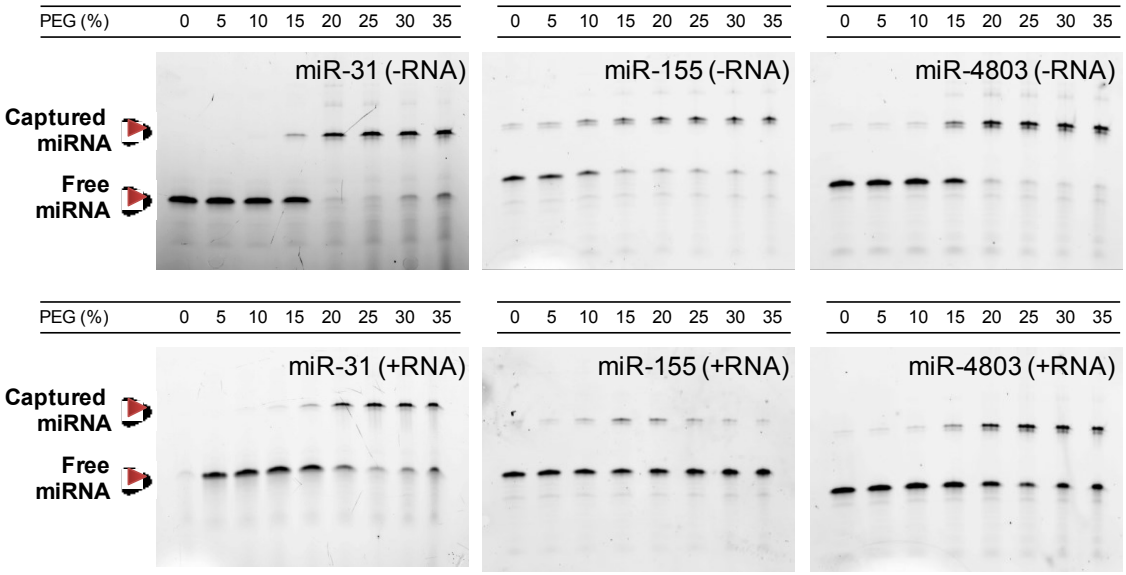

PEG (%)
0 5 10 15 20 25 30 35
miR-31 (-RNA)
miR-155 (-RNA)
miR-4803 (-RNA)
Captured miRNA
Free miRNA
PEG (%)
0 5 10 15 20 25 30 35
miR-31 (+RNA)
miR-155 (+RNA)
miR-4803 (+RNA)
Captured miRNA
Free miRNA

Figure S5.4. Optimization of adapter and enzyme concentration

Raw PAGE images depicting the effects of adapter probe concentration and enzyme amount on the capture efficiency of *miR-31* in idealized buffer conditions and spiked into 500 ng of total RNA. Image analysis was performed on this data to obtain Figure 5.4.

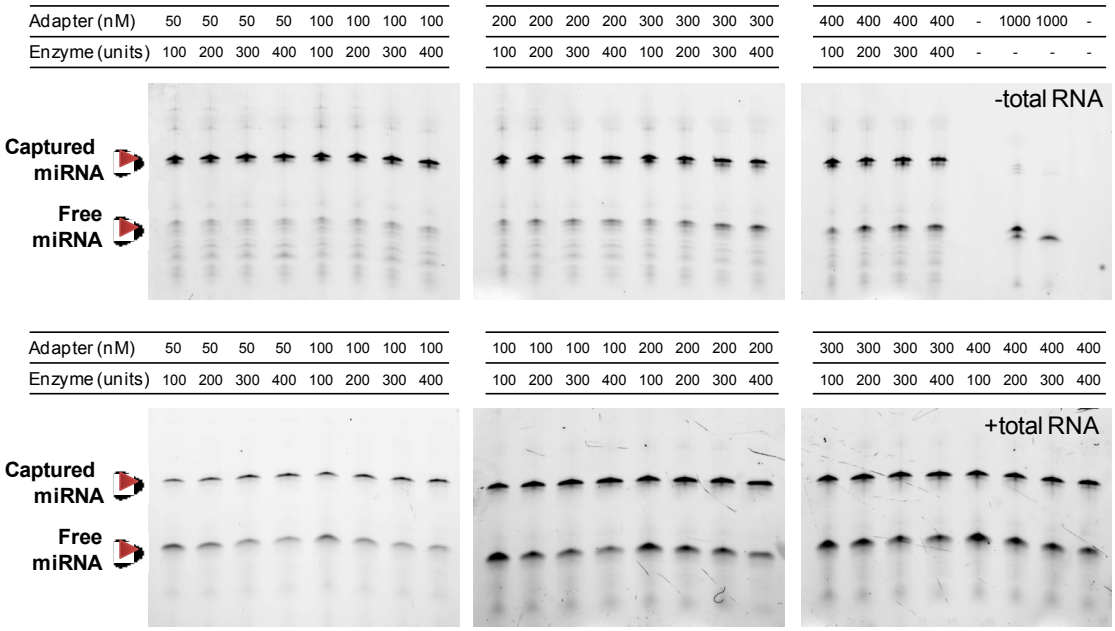

Adapter (nM) 50 50 50 50 100 100 100 100 200 200 200 200 300 300 300 300 400 400 400 400 - 1000 1000 -
Enzyme (units) 100 200 300 400 100 200 300 400 100 200 300 400 100 200 300 400 - - - -
-total RNA
Captured miRNA
Free miRNA
Adapter (nM) 50 50 50 50 100 100 100 100 100 100 100 100 200 200 200 200 300 300 300 300 400 400 400 400
Enzyme (units) 100 200 300 400 100 200 300 400 100 200 300 400 100 200 300 400 100 200 300 400 100 200 300 400
+total RNA
Captured miRNA
Free miRNA

Figure S5.5. Raw PAGE images depicting the effects of incubation time on the capture efficiency of *miR-31* in idealized buffer conditions and spiked into 500 ng of total RNA. Image analysis was performed on this data to obtain Figure 5.5.

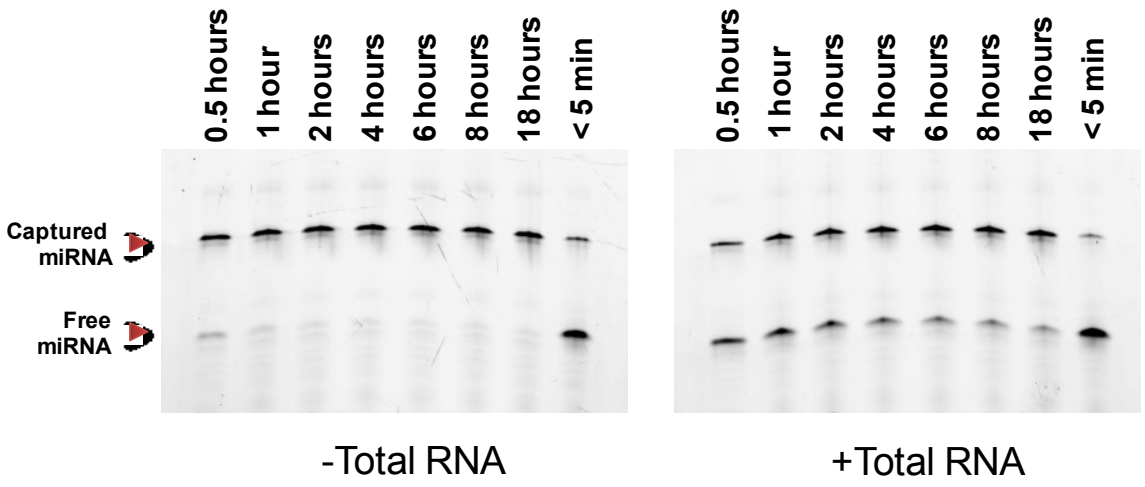

0.5 hours
1 hour
2 hours
4 hours
6 hours
8 hours
18 hours
< 5 min
Captured miRNA
Free miRNA
-Total RNA
+Total RNA

Figure S5.6. Optimization of reaction temperature.

Raw PAGE images depicting the effects of incubation temperature on the capture efficiency and bias across the 20 microRNA panel. Ligation was performed at 4 °C, 25 °C, and 37 °C under identical conditions spiked into 500 ng of total RNA. Image analysis was performed on this data to obtain Figure 5.6.

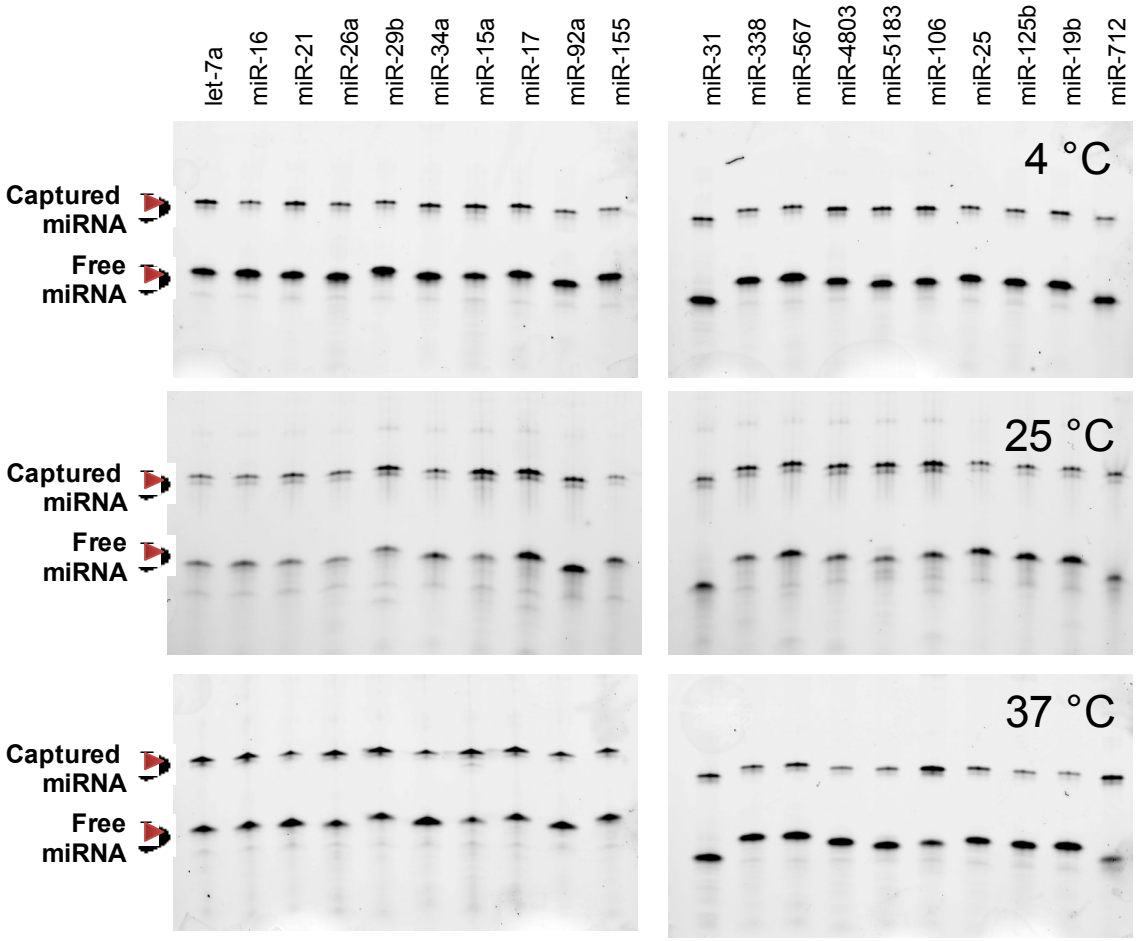

let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106
miR-25
miR-125b
miR-19b
miR-712
Captured miRNA
Free miRNA
4 °C
25 °C
37 °C

Figure S5.7. Comparison of adapter probe designs in microRNA capture.

Raw PAGE images depicting the effects of 4 different adapter probe designs on the capture efficiency and bias across the 20 microRNA panel under spiking conditions. The rA and dA adapter probes are based off of a modified modban probe design with RNA-A vs. DNA-A as the first letter, respectively. The SR1 and SR1-S probes are taken from Zhuang *et al* (39) and Hafner *et al* (38). Image analysis was performed on this data to obtain Figure 5.7.

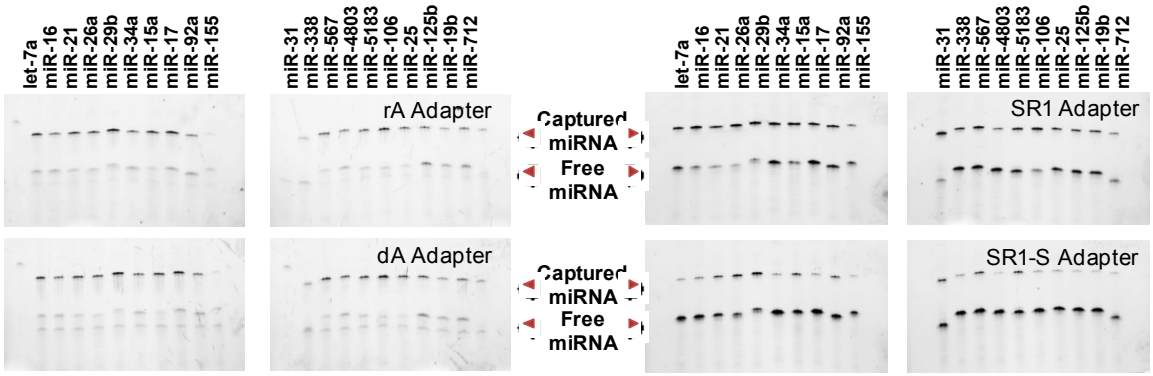
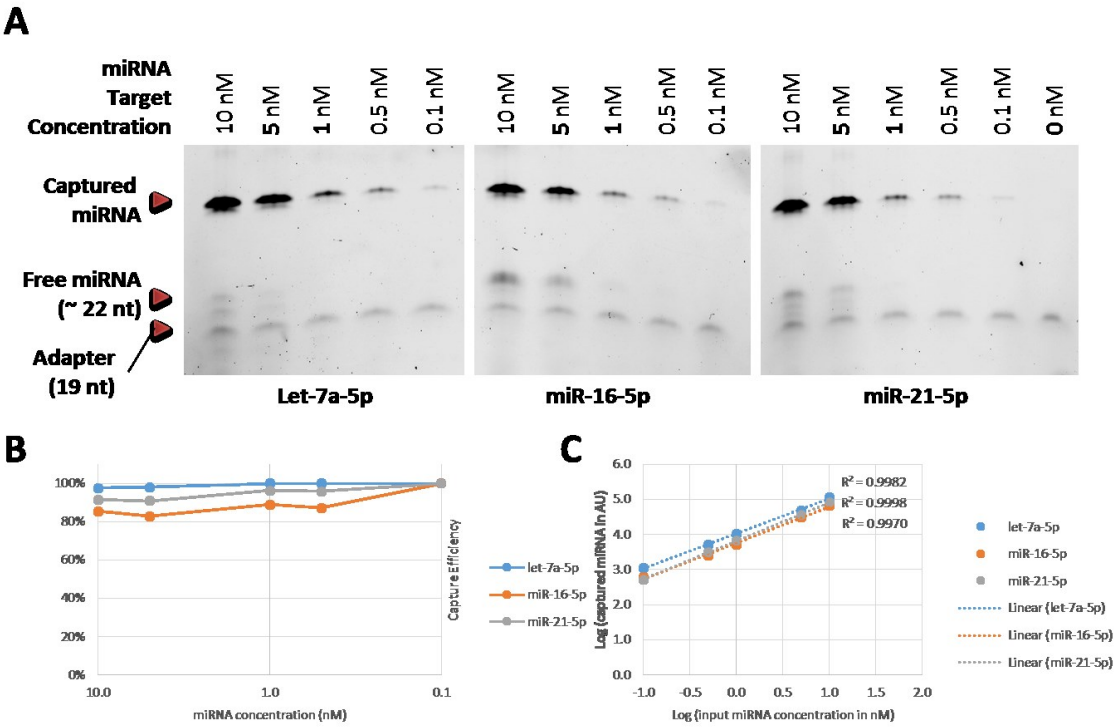

let-7a
miR-16
miR-21
miR-26a
miR-29b
miR-34a
miR-15a
miR-17
miR-92a
miR-155
miR-31
miR-338
miR-567
miR-4803
miR-5183
miR-106
miR-25
miR-125b
miR-19b
miR-712
rA Adapter
dA Adapter
Captured miRNA
Free miRNA
SR1 Adapter
SR1-S Adapter

Figure S5.8. Capturing microRNA within total RNAs.

(A) *Let-7a*, *miR-16*, and *miR-21* were serially diluted from 10 nM - 0.1 nM and captured using the optimized method. The reaction products were then analyzed by denaturing PAGE. (B) The capture efficiency was calculated and plotted against the input miRNA (capture efficiency = captured microRNA / (captured microRNA + free microRNA)). (C) The amount of captured microRNA was plotted against the input microRNA level. Both data indicate that the capture efficiency stays constant across a wide range of input microRNA levels.


A
miRNA Target Concentration
10 nM
5 nM
1 nM
0.5 nM
0.1 nM
0 nM
Captured miRNA
Free miRNA (~ 22 nt)
Adapter (19 nt)
Let-7a-5p
miR-16-5p
miR-21-5p
B
100%
80%
60%
40%
20%
0%
10.0
1.0
0.1
miRNA concentration (nM)
Capture efficiency
let-7a-5p
miR-16-5p
miR-21-5p
C
R² = 0.9982
R² = 0.9998
R² = 0.9970
Log (captured miRNA in AU)
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Log (input miRNA concentration in nM)
let-7a-5p
miR-16-5p
miR-21-5p
Linear (let-7a-5p)
Linear (miR-16-5p)
Linear (miR-21-5p)

Chapter 6: Ligo-miR: Multiplexed MicroRNA Quantification

Song,Y., Kilburn,J.D., Qiu,L., Chen,L., Liu,J.K., Meltzer,J.S. and Wang,T.H., Ligo-miR: a multiplexed ligation assay for microRNA quantification. In Preparation.

6.1. Introduction

Personalized medicine has been a holy grail of medical society since complete of the human genome project. The dream of every individual receiving treatments specifically tailored, based on his/her genetic information, to achieve the best possible outcome has invoked much efforts to develop assays that detect molecular biomarkers – molecules suggesting medical conditions of the patients – from cells and body fluids.

One class of potential biomarkers that are attracting increasing amount of attention is microRNA – the short non-coding RNA that is expressed in a wide variety of tissue cells and organisms. Due to their profound influence in cellular activities such as cell division and tissue differentiation, microRNAs have been suggested to play a key role in cancer progression. Many among over 500 validated human microRNAs have been proven to be abnormally over or under expressed in cancerous tissues or even in body fluids.(134) Such strong link between microRNA expression levels and cancer has fostered expectation of microRNAs serving as valuable cancer

biomarkers. MicroRNAs are also robust; they are resistant to digestion in the body fluid and formalin embedded tissues.(135) Furthermore, microRNAs have been shown to not only distinguish cancerous / non-cancerous states, but also hold the potential to classify cancer types – a favourable trait that current commercially available messenger RNA (mRNA) biomarker tests have yet to demonstrate.(136) Cancer classification is especially powerful since it could not only tell whether the patient needs further medical examination but also narrows down the tissue that deserves a detailed scrutiny.

Encouraged by the great promise, much efforts have been devoted to development of microRNA biomarkers.(137) However, despite substantial investment, no microRNA biomarker has been approved by governments for clinical purposes so far. One of the reasons behind the modest growth is lack of technology that can be used to validate proposed microRNA biomarkers in a quick and cost-effective manner.(138) Microarrays is capable of detecting thousands of microRNAs at a time, hence serving as a crucial tool for wide screening of potential biomarkers. Selected candidates can be studied in detail and on individual basis using Reverse Transcription-quantitative Polymerase Chain Reaction (RT-qPCR). Although sensitive and quantitative, RT-qPCR is costly, tedious, and is typically designed to handle one microRNA in single tube (i.e., single-plex). Hence, the cost of experiment – including

amount of sample, labour, time, and reagents – quickly multiples if the biomarker panel constitutes of multiple microRNAs candidates. Given that thousands of patient samples must be tested under the same condition in biomarker validation processes, RT-qPCR-based methods turn out to be prohibitively expensive and sometimes unrealistic, arresting rapid growth of biomarker industry.

The biogenesis of microRNA also poses a unique hurdle in expression profiling assays. MicroRNAs are first transcribed as long primary microRNAs (pri-microRNAs) that are subsequently processed by Drosha. The resulting precursor microRNAs (pre-microRNAs) are further processed by the Dicer-containing complex, finally resulting in mature microRNAs, which, as part of RNA-induced silencing complex (RISC), regulates gene expression through interacting with mRNA. While functionally active mature microRNAs have distinct ~22 nt sequences and are labelled with 5' phosphate and 3' hydroxyl group, pre-microRNAs have complete mature microRNA sequences but also 3' overhang. Since many microRNA assays including Beadarray(139) and Nanostring(140) detect microRNA based on probe hybridization to the mature microRNA sequences without a clear reaction step selecting against RNAs with the 3' overhang, there is a shortage of multiplexed microRNA profiling assays that yield the true picture of the microRNA expression pattern by confidently distinguishing mature microRNAs from precursors.

To overcome above described problems, we have developed a multiplexed ligation-based assay for microRNA quantification (Ligo-miR), which is capable of quantifying more than 20 mature microRNAs in a robust, sensitive and specific way within a single tube. The assay is simple two-step reaction with minimum requirement of cost and labour, and is specifically designed to detect only mature microRNAs, eliminating complications from the precursors. The assay takes advantages of our previously described high-efficiency, low-bias adapter ligation method(87) and utilizes high specificity of DNA ligase in target recognition. While compatible with a wide range of downstream analysis, the assay is capable of detecting microRNAs of single-digit attomoles or from samples equivalent to 2 ng of total RNA on gel-reading platform. With 20% experiment-to-experiment variability (standard deviation), the assay allows for quantification of samples with 1.2-fold difference in the target amount. Furthermore, we also demonstrate an excellent agreement between the proposed method and the current gold standard RT-qPCR assay. We expect Ligo-miR to play an important role in facilitating the process of microRNA biomarker discovery and validation for cancer diagnosis and prognosis as well as for basic scientific research.

6.2. Materials and Methods

Oligonucleotides. MicroRNA targets and assay probes were synthesized by Integrated DNA Technologies (Coralville, IA, Table 6.1). The adapters were

enzymatically pre-adenylated with T4 RNA ligase. Adapter and Common Probe (CP) are labelled with ddC and Alexa647 at their 3' and 5' termini, respectively, and were purified by HPLC. Discrimination Probes (DPs) and let-7a precursor DNA were labelled with phosphate at their 5' termini and were PAGE-purified. Some of the microRNAs are labelled with Cy3 at their 5' termini and all are PAGE-purified. The sequences of microRNAs were taken from miRBase (www.mirbase.org).

Small RNA Purification from Total RNA. Total RNA extracted from human pancreas tissue was purchased from Life Technologies (Carlsbad, CA). Total RNA extracted from cell lines OE33, SKTG4, QhTRT and ChTRT were generously offered from Dr. Stephen Meltzer. Small RNAs were purified from total RNA using the mirPremier™ microRNA Isolation Kit or Qiagen miRNeasy Kit® combined with Qiagen RNeasy MinElute Kit®, following the manufacturer's instructions.

Ligation Assays. The assay consists of two ligation reactions. In the first step, microRNAs in the sample were captured in 10 uL reaction mixture using our previously described protocol.(87) The second step of the assay, in which CP and DPs were ligated together in presence of target microRNA, was performed by mixing 10 uL of 1st step products and 10 uL of 2nd step reaction master mix. The 2nd step reaction master mix consists of 1X 2nd step buffer (0.1% Triton, 2.5 mM DTT, 0.6 mM ATP), 200 nM CP, 25 nM DP for each

microRNA target, and 40 unit of 9°N™ DNA Ligase (New England Biolabs). The reaction mixture was incubated at 95 °C for 30 seconds and 61 °C for 5 minutes. The incubation was repeated for 50 times.

PAGE Analysis. The samples were analysed on precast 15% TBE-urea polyacrylamide gels (Bio-Rad, Hercules, CA). 10 μL of sample was mixed with 10 μL of loading buffer and heated for 5 minutes at 95°C. The sample was then loaded into the gel and run for 50 min at 300V. The separated gels were scanned using a Typhoon 9410 variable mode imager (GE Healthcare, Piscataway, NJ) and the gel images were analysed using ImageQuant (GE Healthcare, Piscataway, NJ). The fluorescence intensity profile of each lane was extracted and fit with Gaussian curves using Origin (OriginLab, Northampton, MA) to quantify the fluorescence intensity of the bands and their travel distance.

6.3. Results

6.3.1. Assay Principle

The assay consists of two ligation reactions (**Figure 6.1**). MicroRNAs are naturally phosphorylated at 5' terminal. Therefore, if template-independent ligase such as T4 RNA Ligase 1 is added to the target for ligation reaction, a significant amount of microRNA would be circularized or ligated to other RNAs. Such intra-/interconnected molecules are hard to detect, resulting in distortion of the microRNA expression profile. To avoid the problem, a

mutant of T4 RNA Ligase 2 (T4 RNA Ligase 2 truncated, K227Q) is used to ligate an adapter to each microRNA. The enzyme lacks the subunit which adds ATP to (or "adenylates") phosphorylated DNA or RNA; however it can ligate the pre-adenylated DNA or RNA to the 3' terminal of RNA. When optimized, this protocol can add an adapter to the microRNA in a highly efficient manner without significant bias, creating RNA-DNA hybrid.(87) The hybrid molecule serves as the target template in the second step, in which common probe (CP) and discrimination probes (DPs) hybridize to adapter and microRNA sequence of the hybrid, respectively. If CP and DP are perfectly hybridized to the RNA-DNA hybrid, are they ligated together by 9°N™ DNA Ligase and ligation product (C-D complex) is generated. CP is fluorescently labelled and hence so is resulting C-D complexes. The unique length indicates the presence of specific target microRNA while the amount of the specific C-D complex is proportional to the target microRNA in the original sample. The mixture of C-D complex is separated by gel electrophoresis, resulting in distinct gel bands whose fluorescence intensity can be quantified.

6.3.2. Specificity

In order to assess the specificity of the assay, eight of *miR-let-7* family targets (*miR-let7a, b, c, d, e, f, g,* and *i*) were tested on Ligo-miR assay. Each *miR-let-7* family member was individually added to the reaction mixture, which included eight DPs designed to make perfect match with each target. (**Figure**

6.2). For each lane, band intensity of each target was normalized to (divided by) the perfect match band. While slight breed-through of ~15% was observed from let-7a to f, e, and d channel and from let-7b to let-7c channel, unspecific detection was not observed for most of other combinations. Unspecific detection was even less between microRNAs from different families due to larger difference in sequence (**Figure S6.1**).

We also verified that Ligo-miR specifically detects mature microRNA sequences but not precursors (**Figure S6.2**). MicroRNA precursor sequence of *let-7a* and technical control Circ-1 and Circ-2, two artificial sequences, were mixed with mature *let-7a* family targets. DP complementary for mature *let-7* family and Circ-1 and Circ-2 were used in the assay. The assay successfully detected mature microRNA and Circ-1 and Circ-2 without complication from precursor sequence.

6.3.3. Purification of Small RNAs

Ligo-miR was verified to accept both total RNA and small RNA that is purified from total RNA (**Figure S6.3**). As PAGE analysis suggest, large RNAs, including mRNAs among others, can be effectively eliminated by purification columns while minimizing loss of small RNAs (Figure S6.3A and B). To verify that Ligo-miR is compatible with both total RNA and purified small RNA, pancreas total RNA and small RNA purified from the same batch of sample were tested with 24-plex Ligo-miR. The fluorescence intensity of

each detected band was plotted against each other (Figure S6.3C). The fluorescence intensities of each band from both samples made no significant difference. Since purification step can also be used as concentration step if less volume is added to elute the sample than the original sample volume, we included small RNA purification step to the standardized Ligo-miR assay protocol for later experiments.

6.3.4. Sensitivity

The assay sensitivity was assessed first using synthetic target (**Figure 6.3**A). 13 of the 24 microRNA targets were comfortably detectable at as little as input of 1 attomoles. The fluorescence and target input were linearly proportional to each other, enabling accurate microRNA quantification with average R-square value of above 0.98.

MicroRNAs from small input of total RNA was also detected using the assay (Figure 6.3B). 5 microRNAs were detectable at 24 ng of total RNA input. *miR-92a, miR-26a, miR-17, miR-16, miR-let-7a, miR-22, miR-93* and *miR-25* were detected when 1,000 ng of total RNA was added to the reaction.

6.3.5. The Assay Reproducibility

The assay was repeated to assess the reproducibility (**Figure 6.4**). For this purpose, same batch of small RNA from OE33 cell line was tested on two separate experiments. Plotting 17 detected microRNAs against each other, a

highly linear regression was obtained with R^2~0.998. The CV of the experiment was ~15%, allowing accurate microRNA quantification.

6.3.6. Differential Expression Analysis

The same batch of RNA was analysed using both Ligo-miR and RT-qPCR so as to verify the performance of Ligo-miR assay. MicroRNA expression profile of QhTRT, SKGT4 and ChTRT were normalized with OE33 for both assays and plotted against each other. High agreement was obtained with R^2 value of 0.85 ~ 0.95. Some under-expressed microRNAs that were only detected with qPCR but not Ligo-miR were not included in the differential analysis. The experiment ensures that the microRNA expression profile obtained with Ligo-miR is similar to that obtained with qPCR, which is regarded as gold standard.

6.3.7. Quantification of absolute microRNA copies in cell samples

The titration curve (Figure 6.3A) allows for accurate quantification of microRNA within the sample. Assuming each cell has average total RNA of 15 picogram, it is possible to quantify the number of copies of a specific microRNA expressed in a cell (**Table 6.3**). The lowest and highest copies that Ligo-miR detected from OE33, ChTRT, SKGT4 and QhTRT were 40 and 6,000 copies, respectively. As expected, miR-21 was highly expressed in all the cell lines while RNU44, a house keeping miR, was expressed stably in all cell types.

6.4. Discussions

The proposed assay is distinguished by its sole dedication to microRNA targets but not any other type of biomarker. Most ligation assays proposed so far have been based on probe hybridization on short mature microRNA sequences. Such assays can be also applied to detect target sequences in genetic mutation, methylation or mRNA markers; however, they could suffer from complication of precursor microRNAs that share the same mature microRNA sequences but has extra overhang at the 3' terminal. On the other hand, Ligo-miR has excellent specificity to detect only mature microRNA targets: short specific sequences without extra 3' overhang. Stem-loop RT-qPCR shares the same specificity for mature/precursor microRNA differentiation, hence leading to the excellent agreement in differential expression analysis between Ligo-miR and RT-qPCR.

Ligo-miR is sensitive enough to detect sub-attomoles (1,000,000 ~100,000 copies) of synthetic microRNA targets in a multiplexed manner (Figure 6.3A). For lowly expressed microRNAs that exists at 10 copies per cell, Ligo-miR requires total RNA from ~10,000 cells or 150 ng in total. On the other hand, only 1.5 ng of total RNA is required if the target microRNA is expressed at 1,000 copies per cell. Ligo-miR has detected microRNA targets from 24 ng of total RNA input, roughly corresponding to above mentioned range of target input (Figure 6.3B).

Ligo-miR detected 71% of microRNAs (59 out of 88 microRNAs) that was detected by RT-qPCR. Among undetected microRNas, 14 correspond to those that demonstrated relatively high threshold number in RT-qPCR reactions (Ct > 27), suggesting that such microRNAs are under-expressed in the sample.

The distinctions between Ligo-miR and qPCR are summarized in **Table 6.2**. Improved multiplexing capability demonstrated by Ligo-miR is beneficial not only in term of assay use, but also smaller reagent consumption and less demand of sample amount. Although RT-qPCR requires less sample per assay run, each biomarker must be tested on individual basis; hence a large amount of sample is necessary to test all of 24 microRNAs. This means that although RT-qPCR is gold standard for single biomarker analysis (e.g., single SNP detection), it is not designed for routine testing of panel consisting of multiple biomarkers (mRNA, microRNAs), for which Ligo-miR is an excellent solution.

6.5. Table and Figures

Table 6.1. Oligonucleotides.

Name	Sequence 5' -> 3'
Adapter	5' - /5Phos/ACTGTAGGCACCATCAATC/3ddC/ - 3'
Common Probe (CP)	5' - /5Alex647N/GATTGATGGTGCCTACAGT - 3'
DP-Circ-1	5' - /5Phos/GGGATCACTACTCTTTATACAAACTG - 3'
DP-Circ-2	5' - /5Phos/TAGACATTCCAGGCGGTGCGTCAAACTG - 3'
DP-RNU44	5' - /5Phos/AGTCAGTTAGAGCTAATTAAGACCAAACTG - 3'
DP-cel-miR-39-3p	5' - /5Phos/CAAGCTGATTTACACCCGGTGATAATAAACTG - 3'
DP-RNU6B	5' - /5Phos/AAAAATATGGAACGCTTCACGAATTTGCGTAAGTG - 3'
DP-miR-100-3p	5' - /5Phos/CATACCTATAGATACAAGCTTGAGACGCTAATAAACTG - 3'
DP-miR-224-5p	5' - /5Phos/AACGGAACCACTAGTGACTTGCCAAAGACGCTAATAAACTG - 3'
DP-miR-25-3p	5' - /5Phos/TCAGACCGAGACAAGTGCAATGAACCAAAGACGCTAATAAACTG - 3'
DP-miR-205-5p	5' - /5Phos/CAGACTCCGGTGGAATGAAGGAAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-93-5p	5' - /5Phos/CTACCTGCACGAACAGCACTTTGTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-192-5p	5' - /5Phos/GGCTGTCAATTCATAGGTCAGATCTGCTATAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-106b-5p	5' - /5Phos/ATCTGCACTGTCAGCTAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-22-3p	5' - /5Phos/ACAGTTCTTCAACTGGCAGCTTGACCTTATATCCTCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-345-5p	5' - /5Phos/GAGCCCTGGACTAGGAGTCAGCCGGAAGAATGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-29b-3p	5' - /5Phos/AACACTGATTTCAAATGGTGCTATTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-let-7a-5p	5' - /5Phos/AACTATACAACCTACTACCTCAACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-21-5p	5' - /5Phos/TCAACATCAGTCTGATAAGCTACAACTATACAATGATTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG

	- 3'
DP-miR-155-5p	5' - /5Phos/ACCCCTATCACGATTAGCATTAATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-16-5p	5' - /5Phos/CGCCAATATTTACGTGCTGCTAGGATGGATATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-17-5p	5' - /5Phos/CTACCTGCACTGTAAGCACTTTGAATATACTTGATGCATATTATCTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-15a-5p	5' - /5Phos/CACAAACCATTATGTGCTGCTTCCAACACGAACAATACCATGGATGGATATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-34a-5p	5' - /5Phos/ACAACCAGCTAAGACACTGCTTCCCTGTTCCAACCAACACGAACAATACCATGGATGGATATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-26a-5p	5' - /5Phos/AGCCTATCCTGGATTACAATCTCTTTACGAATTTCCCTGTTCCAACCAACACGAACAATACCATGGATGGATATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
DP-miR-92a-3p	5' - /5Phos/ACAGGCCGGGACAAGTATTCATCATTAATCTTTCTCTTTACGTATTATTCTGTTCCAACCAACACGAACAATACCATGGATGGATATTCTGTGGGCTAGACTTTGCTTCTCTCGAAGACCGTATATCTGCTTTAACAACCAAAGACGCTAATAAACTG - 3'
let-7a	5' - /5Cy3/rUrGrArGrGrUrArGrUrArGrGrUrUrGrUrArUrArGrUrU - 3'
miR-16	5' - /5Cy3/rUrArGrCrArGrCrArCrGrUrArArArUrArUrUrGrGrCrG - 3'
miR-21	5' - /5Cy3/rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA - 3'
miR-26a	5' - /5Cy3/rUrUrCrArArGrUrArArUrCrCrArGrGrArUrArGrGrCrU - 3'
miR-29b	5' - /5Cy3/rUrArGrCrArCrCrArUrUrUrGrArArArUrCrArGrUrGrUrU - 3'
miR-34a	5' - /5Cy3/rUrGrGrCrArGrUrGrUrCrUrUrArGrCrUrGrGrUrUrGrU - 3'
let-7a precursor DNA	5' - /5Phos/TGGGATGAGGTAGTAGGTTGTATAGTTTTAGGGTCACACCCACCACTGGGAGATAACTATACAATCTACTGTCTTTCCTA - 3'
circ-1	5' - rGrUrArUrArArArGrArGrUrArGrUrGrArUrGrGrG - 3'
circ-2	5' - rGrArGrGrGrArGrGrGrGrGrUrGrGrArArUrGrUrGrUrA - 3'
SNORD44	5' - rArGTrGrArGTTrArGrArGrGTrArATTrArArGrArGrG - 3'
cel-miR-39-3p	5' - rUrGrArGrGrGrGrGrUrGrUrArArArUrGrArGrGrUrUrG - 3'
RNU6-2	5' -

	rArArArArATrATrGrGrArArGrGrGTTrGrArGrGrArATTTrGrGrGT - 3'
miR-100-3p	5' - rGrArArGrGrUrUrGrUrArUrGrUrArUrArGrGrUrArUrG - 3'
miR-224-5p	5' - rGrArArGrUrGrArGrUrArGrUrGrGrUrUrGrGrGrUrU - 3'
miR-25-3p	5' - rGrArUrUrGrGrArGrUrUrGrUrGrUrGrGrGrUrGrUrGrA - 3'
miR-205-5p	5' - rUrGrGrUrUrGrArUrUrGrGrArGrGrGrGrArGrUrGrUrG - 3'
miR-93-5p	5' - rGrArArArGrUrGrGrUrGrUrUrGrGrUrGrGrArGrGrUrArG - 3'
miR-192-5p	5' - rGrUrGrArGrGrUrArUrGrArArUrUrGrArGrArGrGrG - 3'
miR-106b-5p	5' - rUrArArArGrUrGrGrUrGrArGrArGrUrGrGrArGrArU - 3'
miR-22-3p	5' - rArArGrGrUrGrGrGrArGrUrUrGrArArGrArArGrUrGrU - 3'
miR-345-5p	5' - rGrGrUrGrArGrUrGrGrUrArGrUrGrGrArGrGrGrGrUrG - 3'
miR-29b-3p	5' - rUrArGrGrArGrGrArUrUrUrGrArArArUrGrArGrUrGrUrU - 3'
let-7a-5p	5' - rUrGrArGrGrUrArGrUrArGrGrUrUrGrUrArUrArGrUrU - 3'
miR-21-5p	5' - rUrArGrGrUrUrArUrGrArGrArGrUrGrArUrGrUrUrGrA - 3'
miR-155-5p	5' - rUrUrArArUrGrGrUrArArUrGrGrUrGrArUrArGrGrGrGrU - 3'
miR-16-5p	5' - rUrArGrGrArGrGrArGrGrUrArArArUrArUrUrGrGrGrG - 3'
miR-17-5p	5' - rGrArArArGrUrGrGrUrUrArGrArGrUrGrGrArGrGrUrArG - 3'
miR-15a-5p	5' - rUrArGrGrArGrGrArGrArUrArArUrGrGrUrUrUrGrUrG - 3'
miR-34a-5p	5' - rUrGrGrGrArGrUrGrUrGrUrUrArGrGrUrGrGrUrUrGrU - 3'
miR-26a-5p	5' - rUrUrGrArArGrUrArArUrGrGrArGrGrArUrArGrGrGrU - 3'
miR-92a-3p	5' - rUrArUrUrGrGrArGrUrUrGrUrGrGrGrGrGrGrGrUrGrU - 3'

Table 6.2. Comparison between Ligo-miR and RT-qPCR.

	Taqman RT-qPCR	Ligo-miR EZ
Assay Format	1-plex **1 microRNA per reaction**	24-plex **24 microRNA per reaction**
Assay Step 1	Applied Biosystems Taqman RT Kit Applied Biosystems Taqman microRNA Assay **22 reactions per sample**	Capture Ligation Master Mix **1 reaction per sample**
Assay Step 2	Bio-rad iQ Supermix Applied Biosystems Taqman microRNA Assay **22 reactions per sample**	Coding Ligation Master Mix **1 reaction per sample**
RNA Needed Per Sample	22 reactions x 100 ng per reaction x 3 = **6.6 µg per sample**	1 reaction x 500 ng per reaction x 3 = **1.5 µg per sample**
4x Sample Reagent Cost	$78 per sample x 4 samples x 3 = **$936 total**	$20 per sample x 4 samples x 3 = **$240 total** (proj. retail cost)
4x Sample Reactions Performed	22 reactions per sample x 4 samples x 3 = **264 reactions total**	1 reaction per sample x 4 samples x 3 = **12 reactions total**
4x Sample Analysis Time	**5.5 h** hands-on prep-time + **9 h** machine time	**0.5 h** hands-on prep-time + **9 h** machine time

Table 6.3. Absolute quantification of microRNA

MicroRNA copies per cell was calculated using titration curve (Figure 6.3A).

	OE33	ChTRT	SKGT4	QhTRT
miR-92a-3p	503	724	615	544
miR-26a-5p	443	832	335	329
miR-34a-5p	0	0	0	39
miR-15a-5p	686	940	628	260
miR-17-5p	633	629	623	500
miR-16-5p	1345	1364	1057	594
miR-155-5p	0	0	0	0
miR-21-5p	6020	4101	3625	1457
let-7a-5p	1688	3630	2435	1279
miR-29b-3p	3041	651	699	439
miR-345-5p	0	0	0	0
miR-22-3p	276	475	547	137
miR-106b-5p	1002	185	227	86
miR-192-5p	0	0	0	0
miR-93-5p	398	79	100	55
miR-205-5p	341	0	0	40
miR-25-3p	471	163	148	60
miR-224-5p	0	0	0	0
miR-100-3p	0	0	0	0
RNU6B	3415	1878	2928	176
cel-miR-39-3p	0	0	0	0
RNU44	2886	2957	2384	1574
crc-2	6055	5230	4863	4098
crc-1	11250	8865	9739	7622

Figure 6.1. Assay principle.

The assay is comprised of two ligation steps. MicroRNAs within the sample are first ligated to pre-adenylated adapter using T4 RNA Ligase 2 truncated K227Q to generate RNA-DNA hybrid, which serves as the target template in the second step. In the second step, common probe (CP) and discrimination probes (DPs), respectively, hybridize to adapter and microRNA sequences of the RNA-DNA hybrid. CP is fluorescently labeled at its 5' terminal. Depending on the number of microRNA that is targeted by the assay, multiple DPs, each having unique recognition sequence for target microRNA and different lengths of Variable Length Coding Sequence (VLCS), are added to the reaction. Only if CP and DP are perfectly hybridized to the RNA-DNA hybrid, are they ligated together by 9°N™ DNA Ligase to generate ligation product (C-D complex) with unique length indicating the presence of the target microRNA. The mixture of C-D complexes are separated by denaturing PAGE, generating barcode patterns under fluorescence gel scanner. The band position and fluorescence intensity indicate the identity and amount of microRNA present in the original sample, respectively.

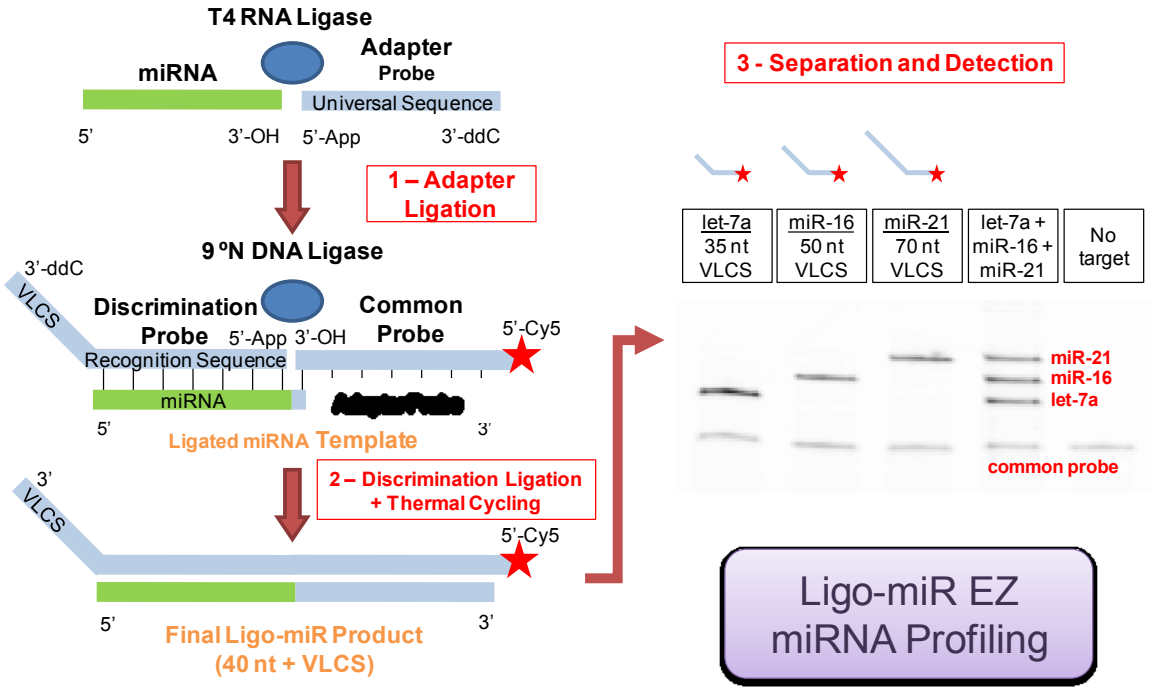
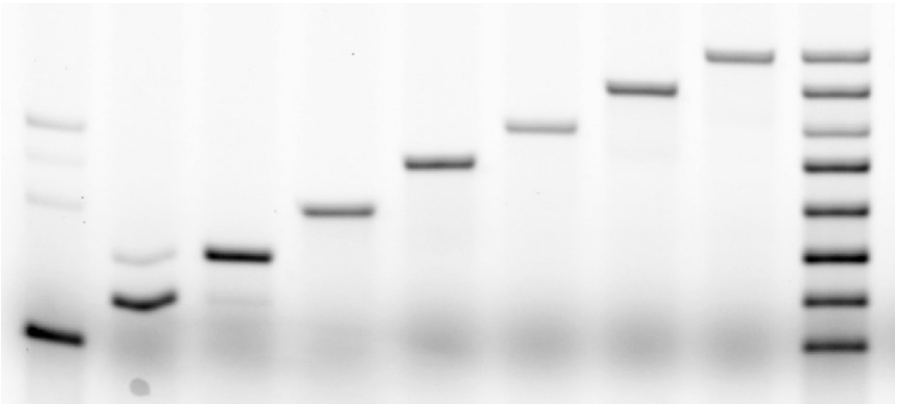

T4 RNA Ligase
Adapter
Probe
miRNA
Universal Sequence
5'
3'-OH
5'-App
3'-ddC
1 – Adapter Ligation
9 °N DNA Ligase
3'-ddC
VLCS
Discrimination Probe
5'-App
3'-OH
Common Probe
5'-Cy5
Recognition Sequence
miRNA
5'
3'
Ligated miRNA Template
2 – Discrimination Ligation + Thermal Cycling
3'
VLCS
5'-Cy5
5'
3'
Final Ligo-miR Product (40 nt + VLCS)
3 - Separation and Detection
let-7a 35 nt VLCS
miR-16 50 nt VLCS
miR-21 70 nt VLCS
let-7a + miR-16 + miR-21
No target
miR-21
miR-16
let-7a
common probe
Ligo-miR EZ miRNA Profiling

Figure 6.2. Specific detection of miR-let7 family.

(A) Each of microRNA *let-7* family (*a, b, c, d, e, f, g,* and *i*) target was tested individually. (B) The bands detected from each lane were normalized to the correct detection. The assay was highly specific with maximum crosstalk of only as much as 16%. (C) Sequences of *let-7* family microRNAs. Sequences differences are highlighted with green on *let-7a* and red on *let-7b* to *let-7i*. *let-7a, d, e*, and *f* share the same sequence at 5' region, leading to highly crosstalk.

(A)



(B)

		Target							
		let-7a	let-7b	let-7c	let-7d	let-7e	let-7f	let-7g	let-7i
Probe	let-7i	-	-	-	-	-	-	-	100%
	let-7g	-	-	-	-	-	-	100%	-
	let-7f	14%	-	-	-	-	100%	-	4%
	let-7e	4%	-	-	-	100%	-	4%	-
	let-7d	9%	-	-	100%	-	-	-	-
	let-7c	-	16%	100%	-	-	-	-	-
	let-7b	-	100%	-	-	-	-	-	-
	let-7a	100%	-	4%	7%	-	-	-	-

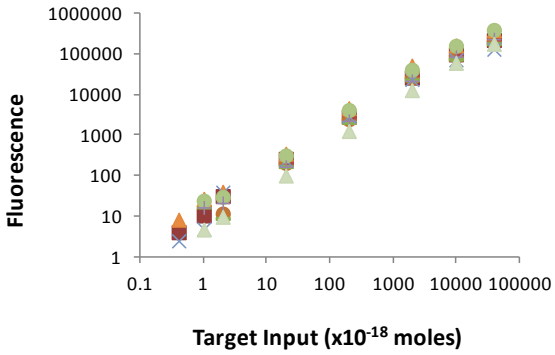
(C)

miRNA	Sequence (5'-3')
let-7a	UGAGGUAGUAGGUUGUAUAGUU
let-7b	UGAGGUAGUAGGUUGUGUGGUU
let-7c	UGAGGUAGUAGGUUGUAUGGUU
let-7d	AGAGGUAGUAGGUUGCAUAGUU
let-7e	UGAGGUAGGAGGUUGUAUAGUU
let-7f	UGAGGUAGUAGAUUGUAUAGUU
let-7g	UGAGGUAGUAGUUUGUACAGUU
let-7i	UGAGGUAGUAGUUUGUGCUGUU

Figure 6.3. Sensitivity assessment.

(A) Synthesized microRNAs varying from 100 to 1 attomoles were added to the reaction. The fluorescence signal of the 24 ligation products were linearly proportional to the amount of input target microRNA. (B) RNA purified from 1,000 ng to 24 ng of total RNA extracted from pancreas were added to the reaction. There was a large variation between the amounts of microRNA detected from the samples.

(A)


1000000
100000
10000
1000
100
10
1
Fluorescence
0.1
1
10
100
1000
10000
100000
Target Input (x10-18 moles)

(B)

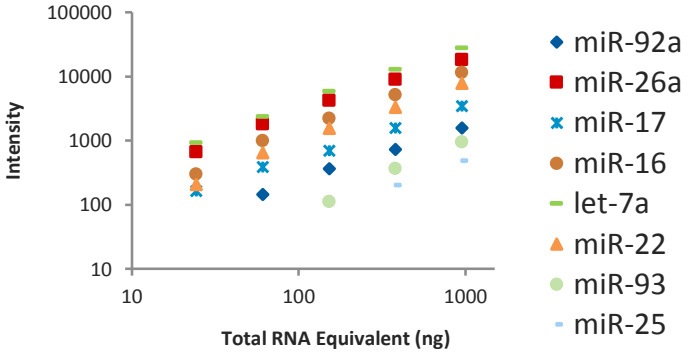

100000
10000
1000
100
10
Intensity
10
100
1000
Total RNA Equivalent (ng)
miR-92a
miR-26a
miR-17
miR-16
let-7a
miR-22
miR-93
miR-25

Figure 6.4. Assay reproducibility.

RNA from OE33 cell line was tested with two separate experiments. Fluorescence intensity of each detected microRNA bands from the two experiments was plotted against each other, resulting in a highly linear relationship. The assay was highly repeatable with correlation coefficient (R^2 value) of 0.998.

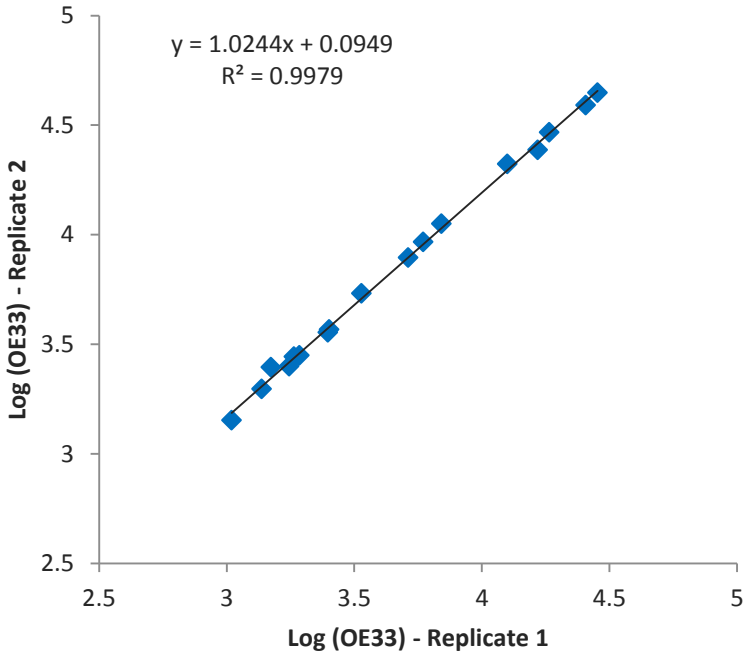
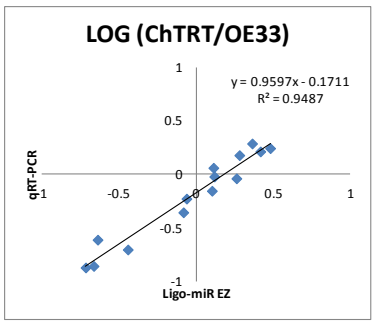
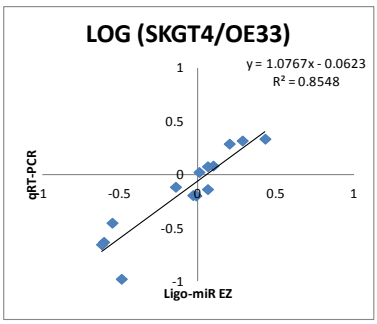
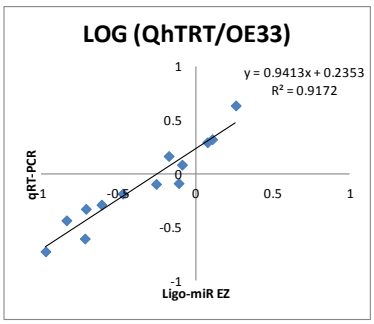

y = 1.0244x + 0.0949
R² = 0.9979
Log (OE33) - Replicate 2
Log (OE33) - Replicate 1
5
4.5
4
3.5
3
2.5
2.5
3
3.5
4
4.5
5

Figure 6.5. Comparison plots of Ligo-miR to RT-qPCR assay.

Log_{10} (microRNA expression level in Cell Line 1 / microRNA expression level in Cell Line 2) was measured by Ligo-miR and RT-qPCR. Signals from cell line ChTRT, SKGT4 and QhTRT were normalized against signals from OE33. No normalization was conducted within individual cell lines. The correlation coefficients was 0.9487, 0.8548, and 0.9172, for ChTRT, SKGT4, and QhTRT , respectively, against OE33.


LOG (ChTRT/OE33)
y = 0.9597x - 0.1711
R² = 0.9487
qRT-PCR
Ligo-miR EZ
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1


LOG (SKGT4/OE33)
y = 1.0767x - 0.0623
R² = 0.8548
qRT-PCR
Ligo-miR EZ
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1


LOG (QhTRT/OE33)
y = 0.9413x + 0.2353
R² = 0.9172
qRT-PCR
Ligo-miR EZ
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1

6.6. Supplementary Information

Figure S6.1. Cross-family detection of microRNA.

Assessment of the assay specificity for microRNA targets in different families. MiR-34a, let-7a, miR-29b, miR-16, miR-26a, and miR-21 bands were detected by the assay without observable crosstalk.

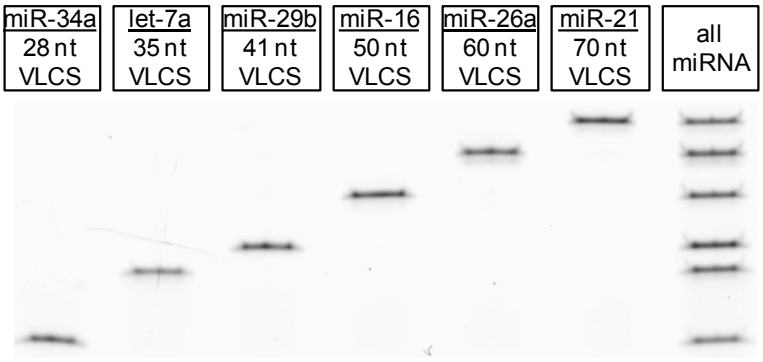

miR-34a
28 nt
VLCS
let-7a
35 nt
VLCS
miR-29b
41 nt
VLCS
miR-16
50 nt
VLCS
miR-26a
60 nt
VLCS
miR-21
70 nt
VLCS
all
miRNA

Figure S6.2. Specific detection of mature microRNA.

let-7a precursor microRNA was spiked as background molecules into a pool of mature *let-7* family microRNA and technical control Circ-1 and Circ-2. Mature *let-7* family microRNAs were added at decreasing concentration (1,000 attomole to 0 attomole) while Circ-1 and Circ-2 as well as *let-7a* precursor were kept at 100 attomole. DP for mature *let-7* family microRNA and Circ-1 and Circ-2 were used. Sub-attomole of mature microRNAs were detected without complication from *let-7a* precursor (not detected due to absence of precursor DP) nor Circ-1 and Circ-2 (always detected at same fluorescence intensity).

miRNA targets (amol)	1000	200	40	8	1.6	.032	.064	0
let-7a precursor (amol)	100	100	100	100	100	100	100	100

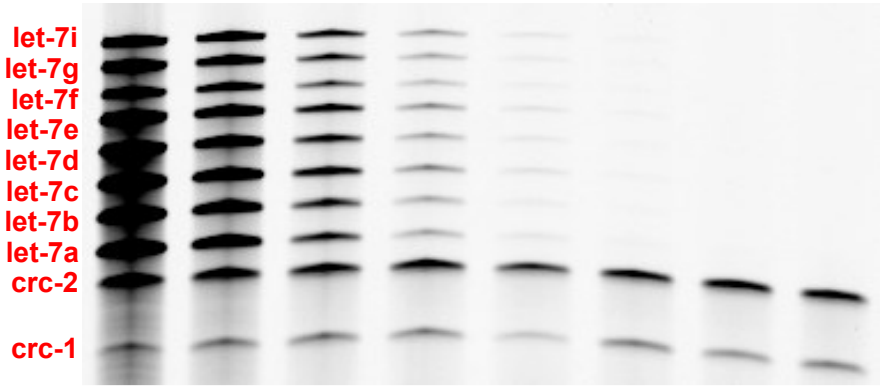
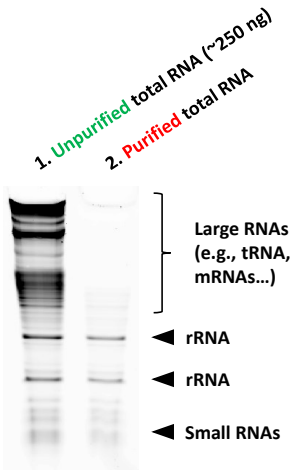
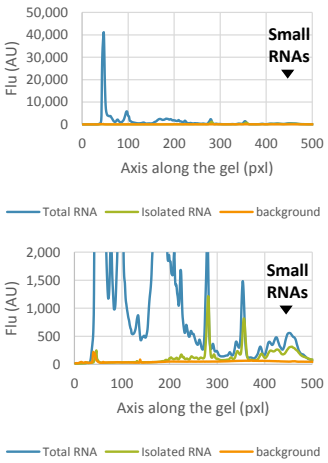


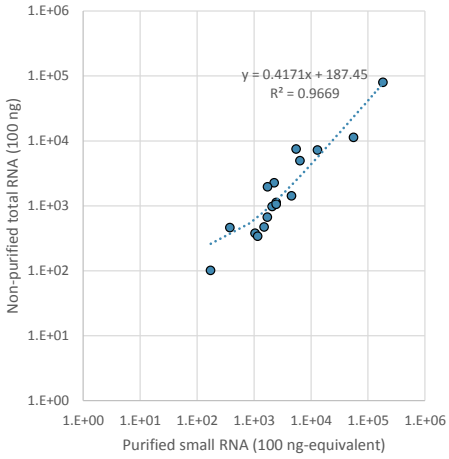
Figure S6.3. Small RNA and total RNA as target for Ligo-miR.

(A) Purification of small RNA from total RNA. Large RNAs are eliminated using the mirPremier™ microRNA Isolation Kit. RNAs are stained with SYBR Safe®. (B) Fluorescence intensity profile of purified (green line) and non-purified (blue line) total RNA. Only small RNAs remain after purification. (C) Result of Ligo-miR assay using purified and non-purified total RNA. The two samples resulted in similar microRNA expression profile with R^2 value of 0.9669.


1. Unpurified total RNA (~250 ng)
2. Purified total RNA
Large RNAs (e.g., tRNA, mRNAs…)
rRNA
rRNA
Small RNAs


Flu (AU)
Small RNAs
Axis along the gel (pxl)
Total RNA
Isolated RNA
background

Flu (AU)
Small RNAs
Axis along the gel (pxl)
Total RNA
Isolated RNA
background


y = 0.4171x + 187.45
R² = 0.9669
Non-purified total RNA (100 ng)
Purified small RNA (100 ng-equivalent)

Chapter 7: Conclusion and Future Directions

In this thesis, I have developed ligation assays for DNA and RNA biomarkers. Ligation technologies not only provide excellent specificity but also are easily incorporated to multiplexed assays, both are essential characteristics for biomarker detection. As more multiplexed assays replace single-plex assays in the future, ligation-based methods discussed in this thesis, Ligo-miR especially, could serve as prototype biomarker detection assays to be used in wide range of research labs or even in clinical settings.

By including universal PCR primer region in DP and CP, all ligation products could be amplified within single tube simultaneously. Combining the benefit of PCR-level sensitivity and multiplexing capability of Ligo-miR, such an assay would be powerful enough to enable multiplexed microRNA detection from smaller amount of RNA input, from body fluids such as serum or plasma, or even from single cells.

The assay can be automated by changing the last PAGE analysis step to capillary electrophoresis or single-molecule free solution hydrodynamic separation (SML-FSHS) invented by our group.(141) Such automation is critical, making Ligo-miR accessible to wider audience. For the assay to be useful as diagnostic/prognostic tool, however, a set panel of microRNA biomarkers needs to be discovered and approved by Food and Drug Administration in USA or equivalent agencies in other countries. However,

as a highly promising biomarker class that potentially elucidates the patient cancer status as well as classifies cancer type, we believe the microRNA biomarker would be continuously investigated by labs all over the world. Ligo-miR could play a great role as a powerful tool kit for their important scientific as well as medical endeavour.

References

1. Stadhouders,R., Pas,S.D., Anber,J., Voermans,J., Mes,T.H. and Schutten,M. (2010) The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *J. Mol. Diagn.,* **12,** 109-117.

2. Sidransky,D., Tokino,T., Hamilton,S.R., Kinzler,K.W., Levin,B., Frost,P. and Vogelstein,B. (1992) Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. *Science,* **256,** 102-105.

3. Sidransky,D. (2002) Emerging molecular markers of cancer. *Nat. Rev. Cancer.,* **2,** 210-219.

4. Hayashi,N., Arakawa,H., Nagase,H., Yanagisawa,A., Kato,Y., Ohta,H., Takano,S., Ogawa,M. and Nakamura,Y. (1994) Genetic diagnosis identifies occult lymph node metastases undetectable by the histopathological method. *Cancer Res.,* **54,** 3853-3856.

5. Brennan,J.A., Mao,L., Hruban,R.H., Boyle,J.O., Eby,Y.J., Koch,W.M., Goodman,S.N. and Sidransky,D. (1995) Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.,* **332,** 429-435.

6. Ho,C.L., Kurman,R.J., Dehari,R., Wang,T.L. and Shih,I. (2004) Mutations of BRAF and KRAS precede the development of ovarian serous borderline tumors. *Cancer Res.,* **64,** 6915-6918.

7. Amado,R.G., Wolf,M., Peeters,M., Van Cutsem,E., Siena,S., Freeman,D.J., Juan,T., Sikorski,R., Suggs,S., Radinsky,R., et al. (2008) Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J. Clin. Oncol.,* **26,** 1626-1634.

8. Massarelli,E., Varella-Garcia,M., Tang,X., Xavier,A.C., Ozburn,N.C., Liu,D.D., Bekele,B.N., Herbst,R.S. and Wistuba,I.I. (2007) KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *Clin. Cancer Res.,* **13,** 2890-2896.

9. van Krieken,J.H., Jung,A., Kirchner,T., Carneiro,F., Seruca,R., Bosman,F.T., Quirke,P., Flejou,J.F., Plato Hansen,T., de Hertogh,G., et al. (2008) KRAS mutation testing for predicting response to anti-EGFR therapy

for colorectal carcinoma: Proposal for an european quality assurance program. *Virchows Arch.,* **453,** 417-431.

10. Syvanen,A.C. (2001) Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.,* **2,** 930-942.

11. Tsuchihashi,Z. and Dracopoli,N.C. (2002) Progress in high throughput SNP genotyping methods. *Pharmacogenomics J.,* **2,** 103-110.

12. Kwok,P.Y. and Chen,X. (2003) Detection of single nucleotide polymorphisms. *Curr. Issues Mol. Biol.,* **5,** 43-60.

13. Kwok,P.Y. (2001) Methods for genotyping single nucleotide polymorphisms. *Annu. Rev. Genomics Hum. Genet.,* **2,** 235-258.

14. Ragoussis,J. (2009) Genotyping technologies for genetic research. *Annu. Rev. Genomics Hum. Genet.,* **10,** 117-133.

15. Lyamichev,V., Mast,A.L., Hall,J.G., Prudent,J.R., Kaiser,M.W., Takova,T., Kwiatkowski,R.W., Sander,T.J., de Arruda,M., Arco,D.A., et al. (1999) Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.,* **17,** 292-296.

16. Lizardi,P.M., Huang,X., Zhu,Z., Bray-Ward,P., Thomas,D.C. and Ward,D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat. Genet.,* **19,** 225-232.

17. Cao,W. (2004) Recent developments in ligase-mediated amplification and detection. *Trends Biotechnol.,* **22,** 38-44.

18. Barany,F. (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci. U. S. A.,* **88,** 189-193.

19. Huang,Y., Zhang,Y.L., Xu,X., Jiang,J.H., Shen,G.L. and Yu,R.Q. (2009) Highly specific and sensitive electrochemical genotyping via gap ligation reaction and surface hybridization detection. *J. Am. Chem. Soc.,* **131,** 2478-2480.

20. Huh,Y.S., Lowe,A.J., Strickland,A.D., Batt,C.A. and Erickson,D. (2009) Surface-enhanced raman scattering based ligase detection reaction. *J. Am. Chem. Soc.,* **131,** 2208-2213.

21. Chen,X., Ying,A. and Gao,Z. (2012) Highly sensitive and selective colorimetric genotyping of single-nucleotide polymorphisms based on enzyme-amplified ligation on magnetic beads. *Biosens. Bioelectron.,* **36,** 89-94.

22. Day,D.J., Speiser,P.W., White,P.C. and Barany,F. (1995) Detection of steroid 21-hydroxylase alleles using gene-specific PCR and a multiplexed ligation detection reaction. *Genomics,* **29,** 152-162.

23. Khanna,M., Park,P., Zirvi,M., Cao,W., Picon,A., Day,J., Paty,P. and Barany,F. (1999) Multiplex PCR/LDR for detection of K-ras mutations in primary colon tumors. *Oncogene,* **18,** 27-38.

24. Chen,X., Livak,K.J. and Kwok,P.Y. (1998) A homogeneous, ligase-mediated DNA diagnostic test. *Genome Res.,* **8,** 549-556.

25. Landegren,U., Kaiser,R., Sanders,J. and Hood,L. (1988) A ligase-mediated gene detection technique. *Science,* **241,** 1077-1080.

26. Cheng,Y., Du,Q., Wang,L., Jia,H. and Li,Z. (2012) Fluorescently cationic conjugated polymer as an indicator of ligase chain reaction for sensitive and homogeneous detection of single nucleotide polymorphism. *Anal. Chem.,* **84,** 3739-3744.

27. Shen,W., Deng,H. and Gao,Z. (2012) Gold nanoparticle-enabled real-time ligation chain reaction for ultrasensitive detection of DNA. *J. Am. Chem. Soc.,* **134,** 14678-14681.

28. Wee,E.J.H., Shiddiky,M.J.A., Brown,M.A. and Trau,M. (2012) eLCR: Electrochemical detection of single DNA base changes via ligase chain reaction. *Chem. Commun.,* .

29. Barany,F. (1991) The ligase chain reaction in a PCR world. *PCR Methods Appl.,* **1,** 5-16.

30. Demchinskaya,A.V., Shilov,I.A., Karyagina,A.S., Lunin,V.G., Sergienko,O.V., Voronina,O.L., Leiser,M. and Plobner,L. (2001) A new approach for point mutation detection based on a ligase chain reaction. *J. Biochem. Biophys. Methods,* **50,** 79-89.

31. Abravaya,K., Carrino,J.J., Muldoon,S. and Lee,H.H. (1995) Detection of point mutations with a modified ligase chain reaction (gap-LCR). *Nucleic Acids Res.,* **23,** 675-682.

32. Yi,P., Jiang,H., Li,L., Dai,F., Zheng,Y., Han,J., Chen,Z. and Guo,J. (2012) A new genotyping method for detecting low abundance single nucleotide mutations based on gap ligase chain reaction and quantitative PCR assay. *Cell Biochem. Biophys.,* **62,** 161-167.

33. Grossman,P.D., Bloch,W., Brinson,E., Chang,C.C., Eggerding,F.A., Fung,S., Iovannisci,D.M., Woo,S. and Winn-Deen,E.S. (1994) High-density multiplex detection of nucleic acid sequences: Oligonucleotide ligation assay and sequence-coded separation. *Nucleic Acids Res.,* **22,** 4527-4534.

34. Iannone,M.A., Taylor,J.D., Chen,J., Li,M.S., Rivers,P., Slentz-Kesler,K.A. and Weiner,M.P. (2000) Multiplexed single nucleotide polymorphism genotyping by oligonucleotide ligation and flow cytometry. *Cytometry,* **39,** 131-140.

35. Jou,C., Rhoads,J., Bouma,S., Ching,S., Hoijer,J., Schroeder-Poliak,P., Zaun,P., Smith,S., Richards,S., Caskey,C.T., et al. (1995) Deletion detection in the dystrophin gene by multiplex gap ligase chain reaction and immunochromatographic strip technology. *Hum. Mutat.,* **5,** 86-93.

36. Livak,K.J. (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet. Anal.,* **14,** 143-149.

37. Baelum,J. and Jacobsen,C.S. (2009) TaqMan probe-based real-time PCR assay for detection and discrimination of class I, II, and III tfdA genes in soils treated with phenoxy acid herbicides. *Appl. Environ. Microbiol.,* **75,** 2969-2972.

38. Li,Y., Guessous,F., Zhang,Y., Dipierro,C., Kefas,B., Johnson,E., Marcinkiewicz,L., Jiang,J., Yang,Y., Schmittgen,T.D., et al. (2009) MicroRNA-34a inhibits glioblastoma growth by targeting multiple oncogenes. *Cancer Res.,* **69,** 7569-7576.

39. Adams,J.S., Ren,S., Liu,P.T., Chun,R.F., Lagishetty,V., Gombart,A.F., Borregaard,N., Modlin,R.L. and Hewison,M. (2009) Vitamin d-directed rheostatic regulation of monocyte antibacterial responses. *J. Immunol.,* **182,** 4289-4295.

40. Tyagi,S. and Kramer,F.R. (1996) Molecular beacons: Probes that fluoresce upon hybridization. *Nat. Biotechnol.,* **14,** 303-308.

41. Zhao,Y., Park,S., Kreiswirth,B.N., Ginocchio,C.C., Veyret,R., Laayoun,A., Troesch,A. and Perlin,D.S. (2009) Rapid real-time nucleic acid sequence-

based amplification-molecular beacon platform to detect fungal and bacterial bloodstream infections. *J. Clin. Microbiol.,* **47,** 2067-2078.

42. Foy,C.A. and Parkes,H.C. (2001) Emerging homogeneous DNA-based technologies in the clinical laboratory. *Clin. Chem.,* **47,** 990-1000.

43. Yahiatene,I., Doose,S., Huser,T. and Sauer,M. (2012) Correlation-matrix analysis of two-color coincidence events in single-molecule fluorescence experiments. *Anal. Chem.,* **84,** 2729-2736.

44. Orte,A., Clarke,R.W. and Klenerman,D. (2011) Single-molecule fluorescence coincidence spectroscopy and its application to resonance energy transfer. *Chemphyschem,* **12,** 491-499.

45. Orte,A., Birkett,N.R., Clarke,R.W., Devlin,G.L., Dobson,C.M. and Klenerman,D. (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U. S. A.,* **105,** 14424-14429.

46. Orte,A., Clarke,R.W. and Klenerman,D. (2008) Fluorescence coincidence spectroscopy for single-molecule fluorescence resonance energy-transfer measurements. *Anal. Chem.,* **80,** 8389-8397.

47. Lymperopoulos,K., Crawford,R., Torella,J.P., Heilemann,M., Hwang,L.C., Holden,S.J. and Kapanidis,A.N. (2010) Single-molecule DNA biosensors for protein and ligand detection. *Angew. Chem. Int. Ed Engl.,* **49,** 1316-1320.

48. Zhang,C.Y. and Yang,K. (2010) Accurate detection of on-state quantum dot and biomolecules in a microfluidic flow with single-molecule two-color coincidence detection. *Anal. Bioanal Chem.,* **397,** 703-708.

49. Zhang,C.Y. and Hu,J. (2010) Single quantum dot-based nanosensor for multiple DNA detection. *Anal. Chem.,* **82,** 1921-1927.

50. Zhang,C.Y., Yeh,H.C., Kuroki,M.T. and Wang,T.H. (2005) Single-quantum-dot-based DNA nanosensor. *Nat. Mater.,* **4,** 826-831.

51. Ranasinghe,R.T. and Brown,T. (2011) Ultrasensitive fluorescence-based methods for nucleic acid detection: Towards amplification-free genetic analysis. *Chem. Commun. (Camb),* **47,** 3717-3735.

52. Yeh,H.C., Ho,Y.P., Shih,I. and Wang,T.H. (2006) Homogeneous point mutation detection by quantum dot-mediated two-color fluorescence coincidence analysis. *Nucleic Acids Res.,* **34,** e35.

53. Ho,Y.P., Kung,M.C., Yang,S. and Wang,T.H. (2005) Multiplexed hybridization detection with multicolor colocalization of quantum dot nanoprobes. *Nano Lett.,* **5,** 1693-1697.

54. Castro,A. and Williams,J.G. (1997) Single-molecule detection of specific nucleic acid sequences in unamplified genomic DNA. *Anal. Chem.,* **69,** 3915-3920.

55. Yeh,H.C., Ho,Y.P. and Wang,T.H. (2005) Quantum dot-mediated biosensing assays for specific nucleic acid detection. *Nanomedicine,* **1,** 115-121.

56. Li,H., Ying,L., Green,J.J., Balasubramanian,S. and Klenerman,D. (2003) Ultrasensitive coincidence fluorescence detection of single DNA molecules. *Anal. Chem.,* **75,** 1664-1670.

57. Cipriany,B.R., Zhao,R., Murphy,P.J., Levy,S.L., Tan,C.P., Craighead,H.G. and Soloway,P.D. (2010) Single molecule epigenetic analysis in a nanofluidic channel. *Anal. Chem.,* **82,** 2480-2487.

58. Neely,L.A., Patel,S., Garver,J., Gallo,M., Hackett,M., McLaughlin,S., Nadel,M., Harris,J., Gullans,S. and Rooke,J. (2006) A single-molecule method for the quantitation of microRNA gene expression. *Nat. Methods,* **3,** 41-46.

59. Schwille,P., Meyer-Almes,F.J. and Rigler,R. (1997) Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys. J.,* **72,** 1878-1886.

60. Ming,T., Chen,H., Jiang,R., Li,Q. and Wang,J. (2011) Plasmon-controlled fluorescence: Beyond the intensity enhancement. *Journal of Physical Chemistry Letters,* **3,** 191.

61. Medintz,I.L., Uyeda,H.T., Goldman,E.R. and Mattoussi,H. (2005) Quantum dot bioconjugates for imaging, labelling and sensing. *Nat. Mater.,* **4,** 435-446.

62. Mackowski,S., Wormke,S., Maier,A.J., Brotosudarmo,T.H., Harutyunyan,H., Hartschuh,A., Govorov,A.O., Scheer,H. and Brauchle,C.

(2008) Metal-enhanced fluorescence of chlorophylls in single light-harvesting complexes. *Nano Lett.,* **8,** 558-564.

63. Kuhn,S., Hakanson,U., Rogobete,L. and Sandoghdar,V. (2006) Enhancement of single-molecule fluorescence using a gold nanoparticle as an optical nanoantenna. *Phys. Rev. Lett.,* **97,** 017402.

64. Muskens,O.L., Giannini,V., Sanchez-Gil,J.A. and Gomez Rivas,J. (2007) Strong enhancement of the radiative decay rate of emitters by single plasmonic nanoantennas. *Nano Lett.,* **7,** 2871-2875.

65. Zhang,J., Fu,Y., Chowdhury,M.H. and Lakowicz,J.R. (2007) Metal-enhanced single-molecule fluorescence on silver particle monomer and dimer: Coupling effect between metal particles. *Nano Lett.,* **7,** 2101-2107.

66. Zin,M.T., Leong,K., Wong,N.Y., Ma,H., Sarikaya,M. and Jen,A.K. (2009) Surface-plasmon-enhanced fluorescence from periodic quantum dot arrays through distance control using biomolecular linkers. *Nanotechnology,* **20,** 015305-4484/20/1/015305. Epub 2008 Dec 5.

67. Akimov,A.V., Mukherjee,A., Yu,C.L., Chang,D.E., Zibrov,A.S., Hemmer,P.R., Park,H. and Lukin,M.D. (2007) Generation of single optical plasmons in metallic nanowires coupled to quantum dots. *Nature,* **450,** 402-406.

68. LeBlanc,S.J., McClanahan,M.R., Jones,M. and Moyer,P.J. (2013) Enhancement of multiphoton emission from single CdSe quantum dots coupled to gold films. *Nano Lett.,* **13,** 1662-1669.

69. Song,J.H., Atay,T., Shi,S., Urabe,H. and Nurmikko,A.V. (2005) Large enhancement of fluorescence efficiency from CdSe/ZnS quantum dots induced by resonant coupling to spatially controlled surface plasmons. *Nano Lett.,* **5,** 1557-1561.

70. Hwang,E., Smolyaninov,I.I. and Davis,C.C. (2010) Surface plasmon polariton enhanced fluorescence from quantum dots on nanostructured metal surfaces. *Nano Lett.,* **10,** 813-820.

71. Fu,Y., Zhang,J. and Lakowicz,J.R. (2009) Silver-enhanced fluorescence emission of single quantum dot nanocomposites. *Chem. Commun. (Camb),* **(3):313-5. doi,** 313-315.

72. Chan,W.C. and Nie,S. (1998) Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science,* **281,** 2016-2018.

73. Zhang,C.Y., Yeh,H.C., Kuroki,M.T. and Wang,T.H. (2005) Single-quantum-dot-based DNA nanosensor. *Nat. Mater.,* **4,** 826-831.

74. Ho,Y.P., Kung,M.C., Yang,S. and Wang,T.H. (2005) Multiplexed hybridization detection with multicolor colocalization of quantum dot nanoprobes. *Nano Lett.,* **5,** 1693-1697.

75. Michalet,X., Pinaud,F.F., Bentolila,L.A., Tsay,J.M., Doose,S., Li,J.J., Sundaresan,G., Wu,A.M., Gambhir,S.S. and Weiss,S. (2005) Quantum dots for live cells, in vivo imaging, and diagnostics. *Science,* **307,** 538-544.

76. Gao,X., Cui,Y., Levenson,R.M., Chung,L.W. and Nie,S. (2004) In vivo cancer targeting and imaging with semiconductor quantum dots. *Nat. Biotechnol.,* **22,** 969-976.

77. Kinkhabwala,A., Yu,Z., Fan,S., Avlasevich,Y., Mu¨llen,K. and Moerner,W.E. (2009) Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna. *Nature Photonics,* **3,** 654.

78. Bek,A., Jansen,R., Ringler,M., Mayilo,S., Klar,T.A. and Feldmann,J. (2008) Fluorescence enhancement in hot spots of AFM-designed gold nanoparticle sandwiches. *Nano Lett.,* **8,** 485-490.

79. Lessard-Viger,M., Rioux,M., Rainville,L. and Boudreau,D. (2009) FRET enhancement in multilayer core-shell nanoparticles. *Nano Lett.,* **9,** 3066-3071.

80. Zhang,J., Malicka,J., Gryczynski,I. and Lakowicz,J.R. (2005) Surface-enhanced fluorescence of fluorescein-labeled oligonucleotides capped on silver nanoparticles. *J Phys Chem B,* **109,** 7643-7648.

81. Lang,X.Y., Guan,P.F., Zhang,L., Fujita,T. and Chen,M.W. (2010) Size dependence of molecular fluorescence enhancement of nanoporous gold. *Applied Physics Letters,* **96,** 073701.

82. Acuna,G.P., Moller,F.M., Holzmeister,P., Beater,S., Lalkens,B. and Tinnefeld,P. (2012) Fluorescence enhancement at docking sites of DNA-directed self-assembled nanoantennas. *Science,* **338,** 506-510.

83. Lang,X., Qian,L., Guan,P., Zi,J. and Chen,M. (2011) Localized surface plasmon resonance of nanoporous gold. *Appl. Phys. Lett.,* **98,** 093701.

84. Hafner,M., Renwick,N., Brown,M., Mihailovic,A., Holoch,D., Lin,C., Pena,J.T., Nusbaum,J.D., Morozov,P., Ludwig,J., et al. (2011) RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA,* **17,** 1697-1712.

85. Jayaprakash,A.D., Jabado,O., Brown,B.D. and Sachidanandam,R. (2011) Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res.,* **39,** e141.

86. Zhuang,F., Fuchs,R.T., Sun,Z., Zheng,Y. and Robb,G.B. (2012) Structural bias in T4 RNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Res.,* **40,** e54.

87. Song,Y., Liu,K.J. and Wang,T.H. (2014) Elimination of ligation dependent artifacts in T4 RNA ligase to achieve high efficiency and low bias microRNA capture. *PLoS One,* **9,** e94619.

88. Lau,N.C., Lim,L.P., Weinstein,E.G. and Bartel,D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in caenorhabditis elegans. *Science,* **294,** 858-862.

89. Vigneault,F., Sismour,A.M. and Church,G.M. (2008) Efficient microRNA capture and bar-coding via enzymatic oligonucleotide adenylation. *Nat. Methods,* **5,** 777-779.

90. Cranston,J.W., Silber,R., Malathi,V.G. and Hurwitz,J. (1974) Studies on ribonucleic acid ligase. characterization of an adenosine triphosphate-inorganic pyrophosphate exchange reaction and demonstration of an enzyme-adenylate complex with T4 bacteriophage-induced enzyme. *J. Biol. Chem.,* **249,** 7447-7456.

91. Sugino,A., Snoper,T.J. and Cozzarelli,N.R. (1977) Bacteriophage T4 RNA ligase. reaction intermediates and interaction of substrates. *J. Biol. Chem.,* **252,** 1732-1738.

92. Torchia,C., Takagi,Y. and Ho,C.K. (2008) Archaeal RNA ligase is a homodimeric protein that catalyzes intramolecular ligation of single-stranded RNA and DNA. *Nucleic Acids Res.,* **36,** 6218-6227.

93. Zhelkovsky,A.M. and McReynolds,L.A. (2011) Simple and efficient synthesis of 5' pre-adenylated DNA using thermostable RNA ligase. *Nucleic Acids Res.,* **39,** e117.

94. Harrison,B. and Zimmerman,S.B. (1984) Polymer-stimulated ligation: Enhanced ligation of oligo- and polynucleotides by T4 RNA ligase in polymer solutions. *Nucleic Acids Res.,* **12,** 8235-8251.

95. Kloosterman,W.P., Wienholds,E., de Bruijn,E., Kauppinen,S. and Plasterk,R.H. (2006) In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat. Methods,* **3,** 27-29.

96. Mineno,J., Okamoto,S., Ando,T., Sato,M., Chono,H., Izu,H., Takayama,M., Asada,K., Mirochnitchenko,O., Inouye,M., et al. (2006) The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res.,* **34,** 1765-1771.

97. Wienholds,E., Kloosterman,W.P., Miska,E., Alvarez-Saavedra,E., Berezikov,E., de Bruijn,E., Horvitz,H.R., Kauppinen,S. and Plasterk,R.H. (2005) MicroRNA expression in zebrafish embryonic development. *Science,* **309,** 310-311.

98. Tang,F., Hajkova,P., Barton,S.C., Lao,K. and Surani,M.A. (2006) MicroRNA expression profiling of single whole embryonic stem cells. *Nucleic Acids Res.,* **34,** e9.

99. Morin,R.D., O'Connor,M.D., Griffith,M., Kuchenbauer,F., Delaney,A., Prabhu,A.L., Zhao,Y., McDonald,H., Zeng,T., Hirst,M., et al. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.,* **18,** 610-621.

100. Lu,J., Getz,G., Miska,E.A., Alvarez-Saavedra,E., Lamb,J., Peck,D., Sweet-Cordero,A., Ebert,B.L., Mak,R.H., Ferrando,A.A., et al. (2005) MicroRNA expression profiles classify human cancers. *Nature,* **435,** 834-838.

101. Volinia,S., Calin,G.A., Liu,C.G., Ambs,S., Cimmino,A., Petrocca,F., Visone,R., Iorio,M., Roldo,C., Ferracin,M., et al. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U. S. A.,* **103,** 2257-2261.

102. Garzon,R., Marcucci,G. and Croce,C.M. (2010) Targeting microRNAs in cancer: Rationale, strategies and challenges. *Nat. Rev. Drug Discov.,* **9,** 775-789.

103. Meyers,B.C., Souret,F.F., Lu,C. and Green,P.J. (2006) Sweating the small stuff: microRNA discovery in plants. *Curr. Opin. Biotechnol.,* **17,** 139-146.

104. Zhang,B., Pan,X., Cobb,G.P. and Anderson,T.A. (2006) Plant microRNA: A small regulatory molecule with big impact. *Dev. Biol.,* **289,** 3-16.

105. Voinnet,O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell,* **136,** 669-687.

106. Keller,A., Leidinger,P., Bauer,A., Elsharawy,A., Haas,J., Backes,C., Wendschlag,A., Giese,N., Tjaden,C., Ott,K., et al. (2011) Toward the blood-borne miRNome of human diseases. *Nat. Methods,* **8,** 841-843.

107. Mitchell,P.S., Parkin,R.K., Kroh,E.M., Fritz,B.R., Wyman,S.K., Pogosova-Agadjanyan,E.L., Peterson,A., Noteboom,J., O'Briant,K.C., Allen,A., et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A.,* **105,** 10513-10518.

108. Valastyan,S., Reinhardt,F., Benaich,N., Calogrias,D., Szasz,A.M., Wang,Z.C., Brock,J.E., Richardson,A.L. and Weinberg,R.A. (2009) A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. *Cell,* **137,** 1032-1046.

109. Chen,C., Ridzon,D.A., Broomer,A.J., Zhou,Z., Lee,D.H., Nguyen,J.T., Barbisin,M., Xu,N.L., Mahuvakar,V.R., Andersen,M.R., et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.,* **33,** e179.

110. Liu,C.G., Calin,G.A., Meloon,B., Gamliel,N., Sevignani,C., Ferracin,M., Dumitru,C.D., Shimizu,M., Zupo,S., Dono,M., et al. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc. Natl. Acad. Sci. U. S. A.,* **101,** 9740-9744.

111. Shingara,J., Keiger,K., Shelton,J., Laosinchai-Wolf,W., Powers,P., Conrad,R., Brown,D. and Labourier,E. (2005) An optimized isolation and labeling platform for accurate microRNA expression profiling. *RNA,* **11,** 1461-1470.

112. Chen,J., Lozach,J., Garcia,E.W., Barnes,B. and Luo,S. (2008) Highly sensitive and specific microRNA expression profiling using BeadArray technology. **36,** e87.

113. Van Nieuwerburgh,F., Soetaert,S., Podshivalova,K., Ay-Lin Wang,E., Schaffer,L., Deforce,D., Salomon,D.R., Head,S.R. and Ordoukhanian,P. (2011) Quantitative bias in illumina TruSeq and a novel post amplification barcoding strategy for multiplexed DNA and small RNA deep sequencing. *PLoS One,* **6,** e26969.

114. Lu,D.P., Read,R.L., Humphreys,D.T., Battah,F.M., Martin,D.I. and Rasko,J.E. (2005) PCR-based expression analysis and identification of microRNAs. *J. RNAi Gene Silencing,* **1,** 44-49.

115. Wang,H., Ach,R.A. and Curry,B. (2007) Direct and sensitive miRNA profiling from low-input total RNA. *RNA,* **13,** 151-159.

116. Barad,O., Meiri,E., Avniel,A., Aharonov,R., Barzilai,A., Bentwich,I., Einav,U., Gilad,S., Hurban,P., Karov,Y., et al. (2004) MicroRNA expression detected by oligonucleotide microarrays: System establishment and expression profiling in human tissues. *Genome Res.,* **14,** 2486-2494.

117. Hafner,M., Landgraf,P., Ludwig,J., Rice,A., Ojo,T., Lin,C., Holoch,D., Lim,C. and Tuschl,T. (2008) Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods,* **44,** 3-12.

118. Thomas,M.F. and Ansel,K.M. (2010) Construction of small RNA cDNA libraries for deep sequencing. *Methods Mol. Biol.,* **667,** 93-111.

119. Ach,R.A., Wang,H. and Curry,B. (2008) Measuring microRNAs: Comparisons of microarray and quantitative PCR measurements, and of different total RNA prep methods. *BMC Biotechnol.,* **8,** 69-6750-8-69.

120. Chen,Y., Gelfond,J.A., McManus,L.M. and Shireman,P.K. (2009) Reproducibility of quantitative RT-PCR array in miRNA expression profiling and comparison with microarray analysis. *BMC Genomics,* **10,** 407-2164-10-407.

121. Git,A., Dvinge,H., Salmon-Divon,M., Osborne,M., Kutter,C., Hadfield,J., Bertone,P. and Caldas,C. (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA,* **16,** 991-1006.

122. Willenbrock,H., Salomon,J., Sokilde,R., Barken,K.B., Hansen,T.N., Nielsen,F.C., Moller,S. and Litman,T. (2009) Quantitative miRNA expression analysis: Comparing microarrays with next-generation sequencing. *RNA,* **15,** 2028-2034.

123. Sorefan,K., Pais,H., Hall,A.E., Kozomara,A., Griffiths-Jones,S., Moulton,V. and Dalmay,T. (2012) Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence,* **3,** 4-907X-3-4.

124. Linsen,S.E., de Wit,E., Janssens,G., Heater,S., Chapman,L., Parkin,R.K., Fritz,B., Wyman,S.K., de Bruijn,E., Voest,E.E., et al. (2009) Limitations and possibilities of small RNA digital gene expression profiling. *Nat. Methods,* **6,** 474-476.

125. Sun,G., Wu,X., Wang,J., Li,H., Li,X., Gao,H., Rossi,J. and Yen,Y. (2011) A bias-reducing strategy in profiling small RNAs using solexa. *RNA,* **17,** 2256-2262.

126. Zhang,Z., Lee,J.E., Riemondy,K., Anderson,E.M. and Yi,R. (2013) High-efficiency RNA cloning enables accurate quantification of miRNA expression by deep sequencing. *Genome Biol.,* **14,** R109.

127. Thomson,J.M., Parker,J., Perou,C.M. and Hammond,S.M. (2004) A custom microarray platform for analysis of microRNA gene expression. *Nat. Methods,* **1,** 47-53.

128. Pak,J. and Fire,A. (2007) Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science,* **315,** 241-244.

129. Viollet,S., Fuchs,R.T., Munafo,D.B., Zhuang,F. and Robb,G.B. (2011) T4 RNA ligase 2 truncated active site mutants: Improved tools for RNA analysis. *BMC Biotechnol.,* **11,** 72-6750-11-72.

130. Yin,S., Ho,C.K. and Shuman,S. (2003) Structure-function analysis of T4 RNA ligase 2. *J. Biol. Chem.,* **278,** 17601-17608.

131. Harrison,B. and Zimmerman,S.B. (1984) Polymer-stimulated ligation: Enhanced ligation of oligo- and polynucleotides by T4 RNA ligase in polymer solutions. *Nucleic Acids Res.,* **12,** 8235-8251.

132. Tessier,D.C., Brousseau,R. and Vernet,T. (1986) Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase. *Anal. Biochem.,* **158,** 171-178.

133. Miyoshi,D. and Sugimoto,N. (2008) Molecular crowding effects on structure and stability of DNA. *Biochimie,* **90,** 1040-1051.

134. Esquela-Kerscher,A. and Slack,F.J. (2006) Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer.,* **6,** 259-269.

135. Mitchell,P.S., Parkin,R.K., Kroh,E.M., Fritz,B.R., Wyman,S.K., Pogosova-Agadjanyan,E.L., Peterson,A., Noteboom,J., O'Briant,K.C., Allen,A., et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A.,* **105,** 10513-10518.

136. Lu,J., Getz,G., Miska,E.A., Alvarez-Saavedra,E., Lamb,J., Peck,D., Sweet-Cordero,A., Ebert,B.L., Mak,R.H., Ferrando,A.A., et al. (2005) MicroRNA expression profiles classify human cancers. *Nature,* **435,** 834-838.

137. Calin,G.A. and Croce,C.M. (2006) MicroRNA signatures in human cancers. *Nat. Rev. Cancer.,* **6,** 857-866.

138. Cho,W.C. (2011) Circulating MicroRNAs as minimally invasive biomarkers for cancer theragnosis and prognosis. *Front. Genet.,* **2,** 7.

139. Chen,J., Lozach,J., Garcia,E.W., Barnes,B., Luo,S., Mikoulitch,I., Zhou,L., Schroth,G. and Fan,J.B. (2008) Highly sensitive and specific microRNA expression profiling using BeadArray technology. *Nucleic Acids Res.,* **36,** e87.

140. Geiss,G.K., Bumgarner,R.E., Birditt,B., Dahl,T., Dowidar,N., Dunaway,D.L., Fell,H.P., Ferree,S., George,R.D., Grogan,T., et al. (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.,* **26,** 317-325.

141. Liu,K.J., Rane,T.D., Zhang,Y. and Wang,T.H. (2011) Single-molecule analysis enables free solution hydrodynamic separation using yoctomole levels of DNA. *J. Am. Chem. Soc.,* **133,** 6898-6901.

Curriculum Vitae

Yunke Song	08/04/14
Name	**Date of this version**

Education History:

Ph. D. expected	2014	Program in Biomedical Engineering Mentor: Dr. Jeff TH Wang	Johns Hopkins School of Medicine
B. S.	2008	Bioscience and Informatics	Keio University, Tokyo

Scholarships:

- Johns Hopkins Institute for Basic Biomedical Sciences Graduate Student Award (Awarded to 8 "promising Ph.D. students" among first year graduate students in the School of Medicine)
- Keio Engineering Society Award (Awarded at commencement to 7 top students among >1000 engineering undergraduate students)
- Heiwa Nakajima Foundation Scholarship (Awarded to 27 promising Japanese students who plan to study in foreign graduate schools)
- Keio University Scholarship (Awarded to 50 top performing students among >1000 engineering undergraduate students)

Peer-Reviewed Journal Publications

1) **Song,Y.**, Kilburn,J.D., Qiu,L., Chen,L., Liu,J.K., Meltzer,J.S. and Wang,T.H., Ligo-miR: a multiplexed ligation assay for microRNA quantification. *In Preparation.*
2) **Song,Y.**, Liu,K.J. and Wang,T.H. Efficient synthesis of stable adenylated DNA and RNA adapters for microRNA capture using T4 RNA Ligase 1. *In Preparation.*
3) **Song,Y.**, Liu,K.J. and Wang,T.H. (2014) Elimination of ligation dependent artifacts in T4 RNA ligase to achieve high efficiency and low bias microRNA capture. *PLoS One*, 9, e94619.
4) Zhang,L.*, **Song,Y.***, Fujita,T., Zhang,Y., Chen,M. and Wang,T.H. (2014) Large enhancement of quantum dot fluorescence by highly scalable nanoporous gold. *Adv Mater*, 26, 1289-1294.
5) **Song,Y.**, Zhang,Y. and Wang,T.H. (2013) Single quantum dot analysis enables multiplexed point mutation detection by gap ligase chain reaction. *Small*, 9, 1096-1105.

(*: These authors contributed equally to this article)

Peer-Reviewed Conference Articles

1) **Song,Y.**, Zhang,L., Chen,M., Wang,T.H., "Single Quantum Dot Fluorescence Enhancement by Tunable Nanoporous Gold," *12th International Conference on Nanotechnology (IEEE Nano 2012)*
2) **Song,Y.**, Zhang,Y., Wang,T.H., "Single Quantum Dot-Based Multiplexed Point Mutation Detection by Gap Ligase Chain Reaction," *15th International Conference on Miniaturized Chemical and Biochemical Analysis Systems (MicroTAS 2011)*
3) Zec,H., Puleo,C., **Song,Y.,** Chen,J., Wang,T.H., "Micro-evaporator as Interconnects to Low-Volume Microfluidic Components," *13th International Conference on Miniaturized Chemical and Biochemical Analysis System (MicroTAS 2009)*

Patent Applications

- PCT App Number PCT/US13/25860, "microRNA Analysis Method", Wang,T.H., Liu,J.K., **Song,Y.**

Service and Leadership

Fall 2012	Teaching assistant for 'Microfabrication Lab'
Fall 2011	Teaching assistant for 'Modeling and Design'